Calc-H02577-59-P32763

THE ACCOUNTING REVIEW

VOLUME 86, No. 4 JULY 2011

26/7/11

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

A FORUM ON INTERNAL CONTROL REPORTING AND CORPORATE DEBT

STEVEN J. KACHELMEIER
Introduction (Editorial)

DAN DHALIWAL, CHRIS HOGAN, ROBERT TREZEVANT, and MICHAEL WILKINS
Internal Control Disclosures, Monitoring, and the Cost of Debt

JEONG-BON KIM, BYRON Y. SONG, and LIANDONG ZHANG
Internal Control Weakness and Bank Loan Contracting: Evidence from SOX Section 404 Disclosures

OTHER ARTICLES

WILLIAM R. BABER, SOK-HYON KANG, and YING LI
Modeling Discretionary Accrual Reversal and the Balance Sheet as an Earnings Management Constraint

JASMIJN C. BOL and STEVEN D. SMITH
Spillover Effects in Subjective Performance Evaluation: Bias and the Asymmetric Influence of Controllability

KENDALL BOWLIN
Risk-Based Auditing, Strategic Prompts, and Auditor Sensitivity to the Strategic Risk of Fraud

FENG CHEN, OLE-KRISTIAN HOPE, QINGYUAN LI, and XIN WANG
Financial Reporting Quality and Investment Efficiency of Private Firms in Emerging Markets

BARBARA MURRAY GREIN and STEFANIE L. TATE
Monitoring by Auditors: The Case of Public Housing Authorities

SHENGQUAN HAO, QINGLU JIN, and GUOCHANG ZHANG
Relative Firm Profitability and Stock Return Sensitivity to Industry-Level News

CHRISTINA A. MASHRUWALA and SHAMIN D. MASHRUWALA
The Pricing of Accruals Quality: January versus the Rest of the Year

DAWN MATSUMOTO, MAARTEN PRONK, and ERIK ROELOFSEN
What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions

JOHN R. ROBINSON, YANFENG XUE, and YONG YU
Determinants of Disclosure Noncompliance and the Effect of the SEC Review: Evidence from the 2006 Mandated Compensation Disclosure Regulations

SIEW HONG TEOH and YINGLEI ZHANG
Data Truncation Bias, Loss Firms, and Accounting Anomalies

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

Board of Directors

President	Kevin D. Stocks, Brigham Young University
President-Elect	Gregory B. Waymire, Emory University
Past President	Nancy A. Bagranoff, Old Dominion University
Vice President	Ira Solomon, University of Illinois at Urbana–Champaign
Vice President–Education	Belverd E. Needles, Jr., DePaul University
Vice President–Education-Elect	Alan Reinstein, Wayne State University
Vice President–Finance	Stacy E. Kovar, Kansas State University
Vice President–International	Kazuo Hiramatsu, Kwansei Gakuin University
Vice President–Professional Relations	Robert H. Colson, Grant Thornton LLP
Vice President–Publications	Jean C. Bedard, Bentley University
Vice President–Publications-Elect	Christine A. Botosan, The University of Utah
Vice President–Research	John Christensen, University of Southern Denmark
Vice President–Sections and Regions	Gail Hoover King, Rockhurst University

The Executive Director of the Association is Tracey E. Sutherland.

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice.

Membership in the Association is available at the following annual rates.

Full membership:

Membership dues of $200.00 include *Accounting Education News*, and the online selection of all three Association-wide journals: *The Accounting Review*, *Accounting Horizons*, and *Issues in Accounting Education*. (This does not include Section journals.) Full members may also opt to receive *hardcopies* of AAA journals for an additional cost:

The Accounting Review	with Print Option $35.00	Total U.S. $235.00
Accounting Horizons	with Print Option $15.00	Total U.S. $215.00
Issues in Accounting Education	with Print Option $15.00	Total U.S. $215.00
All three Association-wide journals	with Print Option $65.00	Total U.S. $265.00

Associate membership:

Full-time students are eligible for associate membership. The print option is not available for Associate members.

All three Association-wide journals	U.S. $30.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Auditing; Diversity; Financial Accounting and Reporting; Forensic and Investigative Accounting; Gender Issues and Worklife Balance; Government and Nonprofit; Information Systems, International Accounting; Management Accounting; Strategic and Emerging Technologies; Teaching, Learning and Curriculum; and Two-Year College. Each Section assesses its own dues, which are collected by the Association.

Editorial correspondence and revisions of manuscripts originally submitted before June 1, 2011 should be sent via email to the incoming Senior Editor, John Harry Evans III, University of Pittsburgh, at tar@katz.pitt.edu • Telephone: (412) 624-8605 • Fax: (412) 648-1693. New manuscripts must be submitted through the Manuscript Submission and Peer Review System, located at http://accr.allentrack.net. The site contains detailed instructions regarding the preparation of files for submission.

The submission fee of $200.00 for members of the AAA or $400.00 for nonmembers may be paid online (VISA or MasterCard only) at https://aaahq.org/AAAforms/journals/tarsubmit.cfm and should be paid before the article is submitted.

Requests for information concerning membership, subscriptions, dues, advertisements, and other matters related to the Association (other than submission of manuscripts) should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Info@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review is indexed in the Social Science Citation Index, which is included in Web of Science.

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rates of $395 for a print copy or at $445 for a print and an electronic copy. The member subscription rate for a hard copy is $35. Electronic access is included with membership dues. Periodicals Postage Paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address changes to ***The Accounting Review***, 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2011 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Steven J. Kachelmeier
Editors
Shannon W. Anderson
J. Harry Evans III
James E. Hunton
Kathryn Kadous
Sanjay G. Kallapur
Ranjani Krishnan
Laureen A. Maines
D. Paul Newman
Thomas C. Omer
Wayne B. Thomas
Mark A. Trombley
Beverly R. Walther
Paul A. Zarowin

59

VOL. 86 JULY 2011 No. 4

A FORUM ON INTERNAL CONTROL REPORTING AND CORPORATE DEBT

STEVEN J. KACHELMEIER
Introduction (Editorial) 1129
DAN DHALIWAL, CHRIS HOGAN, ROBERT TREZEVANT, and MICHAEL WILKINS
Internal Control Disclosures, Monitoring, and the Cost of Debt 1131
JEONG-BON KIM, BYRON Y. SONG, and LIANDONG ZHANG
Internal Control Weakness and Bank Loan Contracting: Evidence from SOX Section 404 Disclosures 1157

OTHER ARTICLES

WILLIAM R. BABER, SOK-HYON KANG, and YING LI
Modeling Discretionary Accrual Reversal and the Balance Sheet as an Earnings Management Constraint 1189
JASMIJN C. BOL and STEVEN D. SMITH
Spillover Effects in Subjective Performance Evaluation: Bias and the Asymmetric Influence of Controllability 1213
KENDALL BOWLIN
Risk-Based Auditing, Strategic Prompts, and Auditor Sensitivity to the Strategic Risk of Fraud 1231
FENG CHEN, OLE-KRISTIAN HOPE, QINGYUAN LI, and XIN WANG
Financial Reporting Quality and Investment Efficiency of Private Firms in Emerging Markets 1255
BARBARA MURRAY GREIN and STEFANIE L. TATE
Monitoring by Auditors: The Case of Public Housing Authorities 1289
SHENGQUAN HAO, QINGLU JIN, and GUOCHANG ZHANG
Relative Firm Profitability and Stock Return Sensitivity to Industry-Level News 1321
CHRISTINA A. MASHRUWALA and SHAMIN D. MASHRUWALA
The Pricing of Accruals Quality: January versus the Rest of the Year 1349
DAWN MATSUMOTO, MAARTEN PRONK, and ERIK ROELOFSEN
What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions 1383

JOHN R. ROBINSON, YANFENG XUE, and YONG YU
Determinants of Disclosure Noncompliance and the Effect of the SEC Review: Evidence from the 2006 Mandated Compensation Disclosure Regulations 1415
SIEW HONG TEOH and YINGLEI ZHANG
Data Truncation Bias, Loss Firms, and Accounting Anomalies 1445

BOOK REVIEWS, Stephen A. Zeff, Editor

Krislert Samphantharak and Robert M. Townsend
Households as Corporate Firms: An Analysis of Household Finance Using Integrated Household Surveys and Corporate Financial Accounting .... SALEHA B. KHUMAWALA 1477
Henri Fortin, Ana Cristina Hirata Barros, and Kit Cutler
Accounting for Growth in Latin America and the Caribbean—Improving Corporate Financial Reporting to Support Regional Economic Development .... ELISEU MARTINS 1480
Stephen A. Zeff
Insights from Accounting History: Selected Writings of Stephen Zeff .... DAVID OLDROYD 1482

Capsule Commentaries .... STEPHEN A. ZEFF 1484

Editorial Policy and Style Information 1487

P 32,763

Editorial Staff of *The Accounting Review*
2010–2011

SENIOR EDITOR
Steven J. Kachelmeier, The University of Texas at Austin

EDITORS
Shannon W. Anderson, University of California, Davis
J. Harry Evans III, University of Pittsburgh
James E. Hunton, Bentley University
Kathryn Kadous, Emory University
Sanjay G. Kallapur, Indian School of Business
Ranjani Krishnan, Michigan State University
Laureen A. Maines, Indiana University
D. Paul Newman, The University of Texas at Austin
Thomas C. Omer, Texas A&M University
Wayne B. Thomas, University of Oklahoma
Mark A. Trombley, The University of Arizona
Beverly R. Walther, Northwestern University
Paul A. Zarowin, New York University

EDITORIAL ASSISTANT
Mary C. Capps, The University of Texas at Austin

RESEARCH ASSISTANT
Tracie Majors, The University of Texas at Austin

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland

PUBLICATIONS DEPARTMENT
Diane Hazard
Lisa Habblitz
Lindsay Wharton Peters
Jan Kovarik (subcontracted)
Copyediting Team (subcontracted)

OTHER AAA STAFF
Barbara Brady
Barbara Gutierrez
Kurt Gardner
Beverly Harrelson
Kathy Casper
Judy Cothern
Man Fong
Pat Stein
Alex Metz
David Boynton
Dierdre Harris
Diane Leger
Michelle Bowles
Roland LaTulip
Gay Williams
Debora Gardner
Peggy Turczyn
Arlene Oliver

Editorial Advisory and Review Board

Anwer S. Ahmed, Texas A&M University
Ashiq Ali, The University of Texas at Dallas
Anil Arya, The Ohio State University
Benjamin C. Ayers, University of Georgia
William R. Baber, Georgetown University
Stephen P. Baginski, University of Georgia
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Linda Smith Bamber, University of Georgia
Orie E. Barron, The Pennsylvania State University
Anne L. Beatty, The Ohio State University
Jean C. Bedard, Bentley University
Robert J. Bloomfield, Cornell University
Jennifer L. Blouin, University of Pennsylvania
Mark T. Bradshaw, Boston College
Lawrence D. Brown, Georgia State University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
Joseph V. Carcello, The University of Tennessee
Mary Ellen Carter, Boston College
Paul K. Chaney, Vanderbilt University
Qiang Cheng, University of Wisconsin–Madison
Michael B. Clement, The University of Texas at Austin
William M. Cready, The University of Texas at Dallas
Antonio Davila, IESE Business School–University of Navarra
Patricia M. Dechow, University of California, Berkeley
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Sunil Dutta, University of California, Berkeley
Leslie G. Eldenburg, The University of Arizona
Patricia M. Fairfield, Georgetown University
David B. Farber, University of Missouri–Columbia
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
James R. Frederickson, Melbourne Business School
Frank B. Gigler, University of Minnesota
Cristi A. Gleason, The University of Iowa
Johnathan C. Glover, Carnegie Mellon University
Steven M. Glover, Brigham Young University
David A. Guenther, University of Oregon
Luzi Hail, University of Pennsylvania
Jeffrey Hales, Georgia Institute of Technology
John R. M. Hand, The University of North Carolina at Chapel Hill
R. Lynn Hannan, Georgia State University
Gilles Hilary, INSEAD
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
Chris E. Hogan, Michigan State University
Ole-Kristian Hope, University of Toronto
Richard W. Houston, The University of Alabama
Raffi J. Indjejikian, University of Michigan
Christopher D. Ittner, University of Pennsylvania
Debra C. Jeter, Vanderbilt University
Karla M. Johnstone, University of Wisconsin–Madison
Bin Ke, Nanyang Technological University

S. Jane Kennedy, University of Washington
Inder K. Khurana, University of Missouri–Columbia
Michael D. Kimbrough, University of Maryland
Ronald R. King, Washington University in St. Louis
April Klein, New York University
W. Robert Knechel, University of Florida
Lisa L. Koonce, The University of Texas at Austin
Susan D. Krische, American University
Jayanthi Krishnan, Temple University
William N. Lanen, University of Michigan
David F. Larcker, Stanford University
Reuven Lehavy, University of Michigan
Carolyn B. Levine, Rutgers, The State University of New Jersey
Robert Libby, Cornell University
Joan L. Luft, Michigan State University
Anne M. Magro, George Mason University
Carol Marquardt, Baruch College–CUNY
Dawn A. Matsumoto, University of Washington
Edward Maydew, The University of North Carolina at Chapel Hill
Brian W. Mayhew, University of Wisconsin–Madison
Sarah E. McVay, University of Utah
Kenneth A. Merchant, University of Southern California
Gregory S. Miller, University of Michigan
Lillian F. Mills, The University of Texas at Austin
Brian Mittendorf, The Ohio State University
Partha S. Mohanram, University of Toronto
Steven J. Monahan, INSEAD
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Arkansas
Venky Nagar, University of Michigan
Terry L. Neal, The University of Tennessee
Karen K. Nelson, Rice University
Edward F. O'Donnell, Southern Illinois University Carbondale
Per M. Olsson, Duke University
Chul W. Park, The University of Hong Kong
Mark E. Peecher, University of Illinois at Urbana–Champaign
Sean A. Peffer, University of Kentucky
Kathy Ruby Petroni, Michigan State University
John D. Phillips, University of Connecticut
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, Stanford University
Marlene A. Plumlee, University of Utah
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shiva Rajgopal, Emory University
Scott Richardson, London Business School
John R. Robinson, The University of Texas at Austin
Darren T. Roulstone, The Ohio State University
Katherine Schipper, Duke University
Karen L. Sedatole, Michigan State University
Catherine M. Shakespeare, University of Michigan
Lakshmanan Shivakumar, London Business School
K. (Shiva) Sivaramakrishnan, University of Houston
Hollis A. Skaife, University of Wisconsin–Madison
Richard G. Sloan, University of California, Berkeley
J. Reed Smith, Indiana University at Indianapolis
Mark T. Soliman, University of Washington
Geoffrey B. Sprinkle, Indiana University
Mary Harris Stanford, Texas Christian University
Phillip C. Stocken, Dartmouth College
K. R. Subramanyam, University of Southern California
Hun-Tong Tan, Nanyang Technological University
Kristy L. Towry, Emory University
Senyo Tse, Texas A&M University
Jennifer Wu Tucker, University of Florida
Irem Tuna, London Business School
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Michael S. Wilkins, Texas A&M University
Richard H. Willis, Vanderbilt University
Arnold M. Wright, Northeastern University
Teri Lombardi Yohn, Indiana University
Xiao-Jun Zhang, University of California, Berkeley
Mark F. Zimbelman, Brigham Young University

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1959–1962	Robert K. Mautz
1962–1965	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William Kinney
1990–1994	A. Rashad Abdel-khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber
2002–2005	Terry Shevlin
2005–2008	Dan S. Dhaliwal

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1129–1130

American Accounting Association
DOI: 10.2308/accr-10046

EDITORIAL

Introduction to a Forum on Internal Control Reporting and Corporate Debt

Steven J. Kachelmeier
The University of Texas at Austin

One of the more surprising things I have learned from my experience as Senior Editor of *The Accounting Review* is just how often a "hot topic" generates multiple submissions that pursue similar research objectives. Though one might view such situations as enhancing the credibility of research findings through the independent efforts of multiple research teams, they often result in unfavorable reactions from reviewers who question the incremental contribution of a subsequent study that does not materially advance the findings already documented in a previous study, even if the two (or more) efforts were initiated independently and pursued more or less concurrently. I understand the reason for a high incremental contribution standard in a top-tier journal that faces capacity constraints and deals with about 500 new submissions per year. Nevertheless, I must admit that I sometimes feel bad writing a rejection letter on a good study, just because some other research team beat the authors to press with similar conclusions documented a few months earlier. Research, it seems, operates in a highly competitive arena.

Fortunately, from time to time, we receive related but still distinct submissions that, in combination, capture synergies (and reviewer support) by viewing a broad research question from different perspectives. The two articles comprising this issue's forum are a classic case in point. Though both studies reach the same basic conclusion that material weaknesses in internal controls over financial reporting result in negative repercussions for the cost of debt financing, Dhaliwal et al. (2011) do so by examining the public market for corporate debt instruments, whereas Kim et al. (2011) examine private debt contracting with financial institutions. These different perspectives enable the two research teams to pursue different secondary analyses, such as Dhaliwal et al.'s examination of the sensitivity of the reported findings to bank monitoring and Kim et al.'s examination of debt covenants.

Both studies also overlap with yet a third recent effort in this arena, recently published in the *Journal of Accounting Research* by Costello and Wittenberg-Moerman (2011). Although the overall "punch line" is similar in all three studies (material internal control weaknesses result in a higher cost of debt), I am intrigued by a "mini-debate" of sorts on the different conclusions reached by Costello and Wittenberg-Moerman (2011) and by Kim et al. (2011) for the effect of material weaknesses on debt covenants. Specifically, Costello and Wittenberg-Moerman (2011, 116) find that "serious, fraud-related weaknesses result in a significant decrease in financial covenants," presumably because banks substitute more direct protections in such instances, whereas Kim et al.

Published Online: July 2011

(2011) assert from their cross-sectional design that company-level material weaknesses are associated with *more* financial covenants in debt contracting.

In reconciling these conflicting findings, Costello and Wittenberg-Moerman (2011, 116) attribute the Kim et al. (2011) result to underlying "differences in more fundamental firm characteristics, such as riskiness and information opacity," given that, cross-sectionally, material weakness firms have a greater number of financial covenants than do non-material weakness firms even *before* the disclosure of the material weakness in internal controls. Kim et al. (2011) counter that they control for risk and opacity characteristics, and that advance leakage of internal control problems could still result in a debt covenant effect due to internal controls rather than underlying firm characteristics. Kim et al. (2011) also report from a supplemental change analysis that, comparing the pre- and post-SOX 404 periods, the number of debt covenants falls for companies both with *and without* material weaknesses in internal controls, raising the question of whether the Costello and Wittenberg-Moerman (2011) finding reflects a reaction to the disclosures or simply a more general trend of a declining number of debt covenants affecting all firms around that time period. I urge readers to take a look at both articles, along with Dhaliwal et al. (2011), and draw their own conclusions. Indeed, I believe that these sorts of debates are healthy for the discipline. A little tension and disagreement helps us to advance.

We do not often get a chance to publish concurrent related studies, but I agree with the reviewers of both Dhaliwal et al. (2011) and Kim et al. (2011) that these studies (and I would include Costello and Wittenberg-Moerman [2011] in the set as well) capture useful synergies that justify the publication of all three in top-tier journals. I trust that our readers will share the same reaction.

REFERENCES

Costello, A. M., and R. Wittenberg-Moerman. 2011. The impact of financial reporting quality on debt contracting: Evidence from internal control weakness reports. *Journal of Accounting Research* 49 (March): 97–136.

Dhaliwal, D., C. Hogan, R. Trezevant, and M. Wilkins. 2011. Internal control disclosures, monitoring, and the cost of debt. *The Accounting Review* 86 (July): 1131–1156.

Kim, J.-B., B. Y. Song, and L. Zhang. 2011. Internal control weakness and bank loan contracting: Evidence from sox section 404 disclosures. *The Accounting Review* 86 (July): 1157–1188.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1131–1156

American Accounting Association
DOI: 10.2308/accr-10043

Internal Control Disclosures, Monitoring, and the Cost of Debt

Dan Dhaliwal
The University of Arizona and Korea University

Chris Hogan
Michigan State University

Robert Trezevant
University of Southern California

Michael Wilkins
Texas A&M University

ABSTRACT: We test the relationship between the change in a firm's cost of debt and the disclosure of a material weakness in an initial Section 404 report. We find that, on average, a firm's credit spread on its publicly traded debt marginally increases if it discloses a material weakness. We also examine the impact of monitoring by credit rating agencies and/or banks on this result and find that the result is more pronounced for firms that are not monitored. Additional analysis indicates that the effect of bank monitoring appears to be the primary driver of these monitoring results. This finding is consistent with the argument that banks are effective delegated monitors for the debt market. The results of this study suggest the need for future research, particularly to test the differential effects of monitoring on the cost of debt compared to the cost of equity.

Keywords: *cost of debt; monitoring of debt; bank monitoring; Section 404 reporting.*

Data Availability: *Data are publicly available from the sources identified in the text.*

I. INTRODUCTION

Section 404 of the Sarbanes-Oxley Act of 2002 (hereafter, Sarbanes-Oxley) requires a publicly traded firm's Form 10-K to contain an audited report concerning the effectiveness of the firm's internal control over financial reporting. Under the Securities and Exchange Commission's (SEC) Final Rule on management's reporting on the effectiveness of internal control

---

We are very grateful to Steven Kachelmeier (the Senior Editor) and two anonymous referees for their valuable and constructive suggestions. We also acknowledge the helpful comments received from workshop participants at The Ohio State University and Monash University.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: April 2009
Accepted: September 2010
Published Online: April 2011

over financial reporting, and under Public Company Accounting Oversight Board (PCAOB) Auditing Standard No. 2 (AS 2), which were in effect during the time period examined in our study, firms are required to disclose in the Section 404 report any material weakness in internal control over financial reporting (SEC 2003; PCAOB 2004). By definition in AS 2, a material weakness indicates a more than remote likelihood that a material misstatement will not be prevented or detected in a firm's financial statements.[1,2] Section 404 became effective for fiscal years ending on or after November 15, 2004, for accelerated filers, defined as U.S. companies with an equity market capitalization exceeding $75 million that file Form 10-K.

In this study, we use the release of an initial Section 404 report as a natural experiment to test whether the cost of a firm's publicly traded debt increases if it discloses a material weakness (hereafter, MW) in internal control over financial reporting. We have several reasons to expect that weak internal control over financial reporting is associated with a higher cost of debt. For example, weak internal control over financial reporting leads to a decrease in the precision of financial reporting numbers. This decrease in precision means that debt investors have less reliable information to assess default risk and to determine compliance with debt covenants (i.e., estimation risk increases), which leads investors to charge a higher cost of debt (Bhojraj and Sengupta 2003). Moreover, weak internal control over financial reporting suggests that managers find it easier to misappropriate a firm's cash flows (Lambert et al. 2007), thus increasing default risk, which again leads investors to charge a higher cost of debt. If weak internal control over financial reporting is associated with a higher cost of debt, then we predict that the disclosure of a MW in an initial Section 404 report is associated with an increase in the cost of a firm's publicly traded debt. We find marginal support for this prediction ($p = 0.099$).

Even though we find that the disclosure of a MW is associated with an increase in a firm's cost of debt, it is possible that this finding may not hold equally between firms that are monitored and firms that are not monitored by credit rating agencies and/or banks. To our knowledge, prior research has not studied the impact of monitoring on the response of markets to the information in a Section 404 report. In Section II, we predict that the increased cost of debt associated with MW disclosure is more pronounced for firms that are not monitored than for firms that are monitored by credit rating agencies and/or banks. We find strong support for this prediction. Additional analysis indicates that the effect of bank monitoring appears to be the primary driver of these monitoring results. This finding is consistent with the argument that banks have a comparative advantage in carrying out information search and monitoring activities.

Ogneva et al. (2007) find that the effect of Section 404 requirements on a firm's cost of equity is strongest for delinquent filers. Also, Moody's Investors Service, Inc. (2005a) states that Moody's views a filing delay as indicating the most serious internal control over financial reporting problems. Consistent with these observations, we find that delinquent filers have an increase in their cost debt after the required initial Section 404 report filing date relative to firms that file a timely report.

---

[1] More specifically, a material weakness is defined as a significant deficiency, or combination of significant deficiencies, that results in more than a remote likelihood that a material misstatement of the annual or interim financial statements will not be prevented or detected (PCAOB 2004). Firms are not required to disclose in a Section 404 report control deficiencies or significant deficiencies that do not rise to the level of a material weakness.

[2] Based on feedback received, including criticism that AS 2 was too costly and resulted in inefficient audits of internal controls, AS 2 was superseded by AS 5 (PCAOB 2007). AS 5 was effective for fiscal years ending on or after November 15, 2007, with early adoption permitted. Like AS 2, AS 5 continues to require disclosures of material weaknesses in internal control over financial reporting. However, the definition of material weakness in AS 5 is somewhat less inclusive than in AS 2. In particular, AS 5 uses the terminology "a reasonable possibility" rather than "a more than remote likelihood" that a material misstatement will not be prevented or detected. We have no reason to believe that our results would have been different under AS 5.

When we partition sample firms by their *ex ante* probability of reporting a MW in an initial Section 404 report, we find that low-probability firms have a larger increase in their cost of debt in response to reporting a MW than high-probability firms. This result is consistent with the notion that the debt market is more surprised when a firm with a low *ex ante* probability of reporting a MW actually reveals a MW than when this occurs for a high *ex ante* probability firm.

Concurrent studies by Kim et al. (2011) and Costello and Wittenberg-Moerman (2011) use bank loan credit spreads to examine some of the same issues that we investigate. The major difference between these two studies and our study is that Kim et al. (2011) and Costello and Wittenberg-Moerman (2011) focus on the impact of Sarbanes-Oxley reporting requirements in the private debt market while we focus on this impact in the public debt market. As further discussed in Section VI, all three studies find that a firm's cost of debt increases after it discloses a MW. Thus, we view the three studies as complementary in that they collectively provide evidence that Sarbanes-Oxley reporting requirements can impact a firm's cost of debt in both the private debt market and the public debt market.

Our study contributes to the literature by examining the effect of Section 404 disclosures on the cost of debt rather than on the cost of equity.[3] One benefit to examining the cost of debt is that measures of the cost of debt are better defined and easier to estimate than measures of the cost of equity.[4] A more important benefit is that examining debt allows us to analyze the effect of monitoring, or a lack of monitoring, by two debt monitors (i.e., credit rating agencies and banks) on the relationship between the cost of debt and the disclosure of a MW. This analysis allows us to assess the impact that monitoring has on the response of markets to the information in a Section 404 report. This assessment of monitoring effects is valuable because, to our knowledge, prior research has not examined this issue. Further, the introduction of monitoring effects into the analysis creates some tension in the prediction that a firm's cost of debt increases if it discloses a MW. Finally, this analysis also provides evidence regarding (1) the argument that banks are effective delegated monitors for the debt market, and (2) the importance of bank monitoring relative to monitoring by credit rating agencies.

A third benefit to examining the cost of debt is that we are able to compare our results to the corresponding cost of equity results in Ashbaugh-Skaife et al. (2009). Based on this analysis, we are able to offer several suggestions for future research. These suggestions involve comparing and contrasting the effect of various monitoring agents on (1) the change in a firm's cost of equity and the disclosure of a MW, and (2) the change in a firm's cost of debt and the disclosure of a MW.

Section II develops the hypotheses. Section III provides a description of the sample. Section IV describes the primary test results. Section V compares these results to the cost of equity results in Ashbaugh-Skaife et al. (2009). Section VI reports the results of additional analyses. Section VII provides a summary of the major findings.

II. HYPOTHESES

Internal Control over Financial Reporting and the Cost of Debt

We have several reasons to expect that weak internal control over financial reporting is associated with a higher cost of debt. For example, cost of debt factors in the likelihood of default. At the same time, weak internal control over financial reporting leads to a reduction in the precision of financial reporting numbers, which decreases the reliability of the information needed to assess the likelihood of default (i.e., estimation risk increases). This decrease in reliability would cause debt investors to charge a higher cost of debt to compensate for their decreased ability to accurately assess the likelihood of default (Bhojraj and Sengupta 2003). Similarly, since debt investors use financial reporting numbers to determine compliance with debt covenants (DeFond and Jiambalvo 1994), this decrease in reliability would lead debt investors to charge a higher cost of debt to compensate for their decreased ability to determine this compliance.

---

[3] Debt is an important source of financing for many firms. For example, among our sample firms, long-term debt as a percentage of equity is, on average, almost 60 percent.

[4] See Botosan and Plumlee (2005) for a discussion of the measurement of the cost of equity, as well as a comparison of different measures of the cost of equity.

Lambert et al. (2007) observe, "[W]hen accounting quality increases, however, managers steal less, so more of the payoff goes to shareholders," where payoff refers to a firm's cash flows. This observation suggests that weak internal control over financial reporting, which decreases accounting quality, provides managers with more opportunities to misappropriate cash flows (i.e., misappropriation risk increases). This increased misappropriation risk increases default risk, thus increasing the cost of debt.

The disclosure of a MW in an initial Section 404 report indicates the potential for a material misstatement in a firm's financial statements due to weak internal control over financial reporting. Thus, based on the preceding arguments, we predict that the disclosure of a MW in an initial Section 404 report is associated, on average, with an increase in a firm's cost of debt. The hypothesis used to test this prediction is:

> **H1:** Firms that initially disclose a MW under Section 404 experience a larger increase in their cost of debt compared to firms with no such disclosure.

Monitoring Effects and the Cost of Debt

Although we predict that a firm's cost of debt increases if it discloses a MW in an initial Section 404 report, it is possible that this average prediction does not hold equally for firms that are monitored compared to firms that are not monitored by credit rating agencies and/or banks. When monitoring debt, a credit rating agency assesses the quality of a firm's internal control over financial reporting. This assessment is reflected in the credit rating assigned to the firm. Consistent with this notion, Moody's Investors Service, Inc. (2005b) indicates that, for firms that report a Section 404 MW, rating actions are not needed in most cases, as the Moody's rating already reflects Moody's impression of the control weaknesses that are eventually disclosed in a Section 404 report.

Based on the preceding, we expect that a firm's credit rating provides the debt market with indirect information about the quality of a firm's internal control over financial reporting, thereby reducing the importance of the information provided to the debt market in a Section 404 report. This reduced importance should lead to a smaller increase in a firm's cost of debt upon the disclosure of a MW than the increase that would occur if no monitoring by credit rating agencies existed. If the preceding is the case, then we expect:

> **H2a:** The increased cost of debt associated with MW disclosure is more pronounced for firms with unrated debt than for firms with rated debt.

Diamond (1984) and Fama (1985), among others, argue that banks serve as effective monitors for the debt market because they have a comparative advantage in carrying out information search and monitoring activities. For example, bankers have an ongoing relationship with a borrower, allowing them access to information about a borrower that is not publicly available (Fama 1985). Also, bankers find it easier to monitor a firm because other providers of corporate debt are more dispersed; thus, monitoring by these other providers suffers from free-rider problems (Diamond 1984). Many empirical studies document support for the notion that banks serve as effective delegated monitors (e.g., James 1987; Datta et al. 2000; Bharath et al. 2008; Altman et al. 2010).

Given the preceding, we argue that firms without bank loans exhibit a larger increase in their cost of debt following the disclosure of a MW in an initial Section 404 report than firms with bank loans. If a bank makes a loan to a given firm, the bank would assess the quality of the firm's internal control over financial reporting; this assessment would be reflected in the interest rate on the loan. In other words, banks would price protect against internal control over financial reporting problems. Further, a firm that has bank loans must disclose information about the terms of these loans in Form 10-K. Also, there are databases that provide detailed information about the terms of a firm's bank loan agreements (e.g., DealScan). This information would provide the debt market with indirect information about the quality of a firm's internal control over financial reporting, thereby reducing the importance of the information provided in a Section 404 report.

Further, the debt market could use this information to mimic the behavior of banks and also price protect against internal control over financial reporting problems, thereby reducing (relative to firms not subject to bank monitoring) the reaction of the cost of debt to the disclosure of a MW.[5] Based on the preceding arguments, we expect:

> **H2b:** The increased cost of debt associated with MW disclosure is more pronounced for firms that are not subject to bank monitoring than for firms that are subject to bank monitoring.

We initially test H2a and H2b separately. We then test these hypotheses jointly to determine if the monitoring effect of credit rating agencies and that of banks each has its own incremental effect on the relationship between the cost of debt and the disclosure of a MW, as opposed to one effect subsuming the other.

Delinquent Filers and the Cost of Debt

Ogneva et al. (2007) find that the effect of Section 404 on the cost of equity is strongest for delinquent filers. Additionally, Moody's Investors Service, Inc. (2005a, 1–2) views a delay in filing a Section 404 report as indicating the most serious internal control over financial reporting problems for the following reasons:

- Late filers may heighten uncertainty about the nature and extent of their control deficiencies while the company and its auditor complete their evaluation of the company's controls. This uncertainty, in turn, raises questions about the reliability of the company's financial data, particularly un-audited data, the ultimate timing of filing reports with the SEC, and the nature of the company's plan to remediate control problems.

Regarding liquidity, later filers may:

- *Temporarily lose access to the public capital markets* because their SEC filings are "delinquent";
- *Also miss the deadline for filing audited financial statements*, which, in turn, could cause uncertainty among market participants and affect the amount and terms of capital available to the company; and
- *Violate covenants in credit agreements and/or indentures.* A common covenant requires timely filing of financial data with the SEC.

Based on the preceding, we test the following hypothesis:

> **H3:** Firms that delay filing an initial Section 404 report experience an increase in their cost of debt relative to timely filers.

---

[5] If the debt market did not behave in this manner, then it is possible that the cost of debt would significantly increase when a firm that a bank is supposed to be monitoring discloses a MW in an initial Section 404 report. However, our results suggest that this is not the case.

The *Ex Ante* Probability of Reporting a MW and the Cost of Debt

Ashbaugh-Skaife et al. (2009) document that the increase in the cost of equity for firms with a low *ex ante* probability of reporting a MW is greater than the increase for high-probability firms when a Section 302 material weakness (hereafter, Section 302 MW) or a Section 404 MW is disclosed.[6] This finding is consistent with the notion that the equity market is more surprised when a firm with a low *ex ante* probability of reporting a MW actually reveals a MW than when this occurs for a high *ex ante* probability firm. Based on this evidence, we expect:

H4: The increased cost of debt associated with MW disclosure is more pronounced for low *ex ante* probability of reporting MW firms than for high *ex ante* probability firms.

III. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

The initial sample of Section 404 disclosures consists of 4,536 firm-years identified as reporting or not reporting a Section 404 MW in their Form 10-K filings between November 3, 2004, and March 3, 2006, according to the Audit Analytics Internal Controls (hereafter, Audit Analytics) database. We eliminate firm-years with erroneous or duplicate CIK codes (these are the firm-identifier codes that Audit Analytics uses), duplicate or triplicate firm-years (e.g., where a firm disclosed internal control status updates in a subsequent year's Form 10-K), and firm-years not in the Compustat database. After applying these screens, we are left with 3,640 (3,087 non-MW and 553 MW) unique firms making an initial Section 404 disclosure. The firms come from a broad range of industries. For example, only four two-digit SIC industries comprise more than 5 percent of the firms, with the greatest concentration in SIC 60 (Depository Institutions, 10.93 percent of firms) and SIC 73 (Business Services, 10.27 percent of firms).

To obtain yield data, we use the National Association of Securities Dealers (NASD) Trade Reporting and Compliance Engine (TRACE) fixed income securities transaction data. All brokers/dealers who are NASD members have an obligation to report transactions in corporate debt instruments via the TRACE system. TRACE should provide a relatively complete picture of trading in corporate debt instruments. For example, by the time of our sample period, (1) "As a general rule, the vast majority of dollar denominated corporate debt instruments must be reported to TRACE" (NASD 2004, 4), and (2) "approximately 99% of all transactions and 95% of par value in the TRACE-eligible securities market are disseminated immediately" (SEC 2005, 5).[7] Further, the data collected and disseminated on these transactions would cover the majority of all transactions in corporate bonds because "[m]ost corporate bonds trade in the over-the-counter (OTC) market ... The OTC market is much larger than the exchange markets, and the vast majority of bond transactions, even those involving exchange-listed issues, take place in this market" (Securities Industry and Financial Markets Association [SIFMA] 2009, 1).

---

[6] Ashbaugh-Skaife et al. (2009) use both Section 302 and Section 404 disclosures in their analysis. Section 302, which was finalized and implemented on August 29, 2002, requires management to certify in quarterly and annual reports that disclosure control processes and procedures have been designed and implemented, and to conclude on the effectiveness of disclosure controls. Both Sections 302 and 404 require the disclosure of material weaknesses in internal control over financial reporting. Under Section 302, the disclosure is the responsibility of management, while under Section 404, the disclosure is the responsibility of the firm's independent auditor, as well as management.

[7] For example, in 2004 TRACE was collecting data on secondary over-the-counter market transactions for approximately 23,000 publicly traded corporate bond issues, including investment-grade debt, high-yield debt, and unrated debt, and on a typical day $20 billion par value of corporate bonds turned over in approximately 25,000 transactions (NASD 2004). In 2005, transactions not reported immediately were subject to a maximum ten-business-day reporting delay, except for trades of Rule 144A securities, which were not disseminated because of the limited transferability and tradability of securities sold pursuant to Rule 144A (SEC 2005).

TRACE provides transaction-specific information, including price and yield. Using this data, we calculate the transaction-specific credit spread as yield to maturity, or YTM, on a given corporate debt instrument transaction minus YTM (obtained from Federal Reserve Board [2006]) on a Treasury security matched on time to maturity.[8]

We remove from the sample firms with missing required data, and four firms for which the absolute value of the test variable *ΔDefaultRisk* exceeds 400 percent.[9] We also remove 43 firms that report a Section 302 MW prior to their initial Section 404 report.[10,11] These procedures yield a sample of 577 (531 non-MW and 46 MW) firms. Table 1 reports descriptive statistics for these firms.

Similar to previous studies, MW firms are smaller, less profitable, and more distressed than non-MW firms.[12] Also, for both MW and non-MW firms, there is an increase in credit spreads in the time period following the initial Section 404 report filing date, indicating that our tests should control for this general increase in credit spreads over the test period. Moreover, MW firms have a higher pre-report credit spread than non-MW firms, suggesting that the debt market, at least partially, can anticipate which firms are more likely to disclose a MW. This last observation indicates that our tests should control for pre-report credit spreads.

Finally, if we partition the *Spread* (*pre*) data in Table 1 by bank monitoring status, we find that, for bank-monitored firms over days [–45, –3] relative to the initial Section 404 report filing date, the mean spread for MW firms in excess of the mean spread for non-MW firms equals 3.22 percent. For firms not subject to bank monitoring, the corresponding difference is 1.69 percent. Finally, consistent with the notion that banks serve as effective monitors and that the debt market mimics bank behavior, the difference between 3.22 percent and 1.69 percent is marginally significant ($p = 0.062$).

IV. TEST DESIGN AND RESULTS

Test of H1: Internal Control Disclosures and the Cost of Debt

Test Design

To examine if firms that initially disclose a Section 404 MW experience an increase in their cost of debt relative to firms with no such disclosure, we estimate Model (1):[13]

---

[8] In these calculations, to eliminate outliers, we delete the top and bottom 1 percent of transaction-specific yield observations.

[9] Our conclusions do not change if we estimate our models after setting *ΔDefaultRisk* for these four firms equal to the next most extreme value of *ΔDefaultRisk* observed in the distribution of *ΔDefaultRisk*.

[10] These firms are removed because (1) the disclosure of a prior Section 302 MW may preempt any information provided by the disclosure of a MW in an initial Section 404 report, and (2) we do not want firms that disclose a prior Section 302 MW, but disclose no MW in an initial Section 404 report, included in the sample because the remediation of a material weakness is not the primary interest of our study.

[11] Approximately 85 percent of our sample firms have complete data on debt seniority. We do not want to reduce the sample size for our regression tests further by requiring data on debt seniority. If we estimate our models after adding the variable *Seniority*, coded 1 for firms that have trades in only senior/senior secured debt instruments, and 0 for firms that either have trades in at least some less senior debt instruments or have missing seniority data, the estimated coefficient on *Seniority* is negative and not significant at conventional levels (most significant $p = 0.115$). The only notable change in our results is that the estimated coefficient on *MW* in Model (1) is positive at $p = 0.120$ (rather than $p = 0.099$). Based on this analysis, we conclude that the seniority of debt does not appear to have an influence on our test results.

[12] The previous studies that we refer to here are Ge and McVay (2005), Ashbaugh-Skaife et al. (2007, 2009), Doyle et al. (2007), Ogneva et al. (2007), Beneish et al. (2008), Hogan and Wilkins (2008), Crabtree et al. (2009), Elbannan (2009), and Kim et al. (2011).

[13] Appendix A summarizes the data sources used to calculate model variables.

TABLE 1

Descriptive Statistics

	531 Non-MW Firms			46 MW Firms		
Variable	**Mean**	**Median**	**Std. Dev.**	**Mean**	**Median**	**Std. Dev.**
Yield (pre)	4.982	4.822	2.230	7.788*	6.042*	6.784
Yield (post)	5.520	5.260	2.641	8.539*	6.529*	6.844
Spread (pre)	1.308	1.011	2.240	4.050*	2.487*	6.750
Spread (post)	1.676	1.311	2.693	4.737*	2.634*	6.811
DefaultRisk (pre)	−1.800	−1.841	2.039	−0.578*	−0.691*	2.045
DefaultRisk (post)	−1.431	−1.652	1.685	−0.415*	−0.248*	2.010
Assets (pre)	7760.93	2878.48	14151.94	4427.13*	1317.92*	7560.75
Assets (post)	7915.96	2978.40	14411.51	4332.26*	1298.93*	7353.87
CFO (pre)	0.054	0.060	0.078	0.016*	0.016*	0.055
CFO (post)	0.086	0.092	0.101	0.037*	0.041*	0.082
Leverage (pre)	0.319	0.271	0.203	0.299	0.290	0.169
Leverage (post)	0.315	0.266	0.208	0.310	0.284	0.171
CFVol (pre)	0.043	0.036	0.031	0.045	0.033	0.037
CFVol (post)	0.045	0.038	0.032	0.044	0.036	0.036
TrsYTM (pre)	3.674	3.754	0.640	3.738	3.900	0.712
TrsYTM (post)	3.844	3.933	0.604	3.802	3.917	0.645
Restructure (pre)	0.311	0	0.463	0.478*	0	0.505
Restructure (post)	0.382	0	0.486	0.478	0	0.505
WriteDown (pre)	0.130	0	0.337	0.152	0	0.363
WriteDown (post)	0.196	0	0.397	0.304	0	0.465
M&A (pre)	0.043	0	0.204	0.022	0	0.147
M&A (post)	0.055	0	0.227	0.022	0	0.147
Restate (pre)	0.122	0	0.327	0.111	0	0.318
Restate (post)	0.127	0	0.334	0.178	0	0.387

---

* Indicates that the "531 Non-MW Firms" value is significantly different from the "46 MW Firms" value at the two-tailed 5 percent level or better, based on a t-test of difference in means or a Wilcoxon two-sample test for medians.

Table 1 provides summary statistics for the 531 sample firms that do not report a MW and the 46 sample firms that do report a MW in an initial Section 404 report.

Variable Definitions:

Yield (pre) = mean (for all TRACE-reported trades in a firm's debt issues) YTM over days [−45, −3] relative to the initial Section 404 report filing date;

Yield (post) = mean (for all TRACE-reported trades in a firm's debt issues) YTM over days [+3, +45];

Spread (pre) = mean (for all TRACE-reported trades in a firm's debt issues) credit spread over days [−45, −3], where the credit spread for each trade is YTM on the relevant debt issue minus YTM on a matched-by-maturity Treasury security;

Spread (post) = mean (for all TRACE-reported trades in a firm's debt issues) credit spread over days [+3, +45];

DefaultRisk = score from the bankruptcy model of Ohlson (1980);

Assets = total assets in $ millions;

CFO = cash flow from operations scaled by total assets;

Leverage = long-term debt scaled by total assets;

CFVol = standard deviation over the last 20 quarters of cash flow from operations scaled by total assets;

For each sample firm *i*, $TrsYTM_i$ *(pre)* = the mean of all "YTM on a matched-by-maturity Treasury security" used to calculate $Spread_i$ over days [−45, −3], and $TrsYTM_i$ *(post)* = the mean of all "YTM on a matched-by-maturity Treasury security" used to calculate $Spread_i$ over days [+3, +45]; and

Restructure (WriteDown, M&A, Restate) = an indicator variable equal to 1 if restructuring (writedown, M&A, restatement) activity occurs in a given period, and 0 otherwise.

(continued on next page)

TABLE 1 (continued)

For *DefaultRisk*, *Assets*, *CFO*, and *Leverage*, (*pre*) refers to value for Q3 and (*post*) refers to value for Q4, where Q4 is the fourth quarter of the fiscal year that is reported on in a firm's initial Section 404 report and Q3 is the preceding quarter. For *CFVol*, (*pre*) refers to the last 20 quarters ending in Q3, and (*post*) refers to the last 20 quarters ending in Q4. For *Restructure*, *WriteDown*, and *M&A*, (*pre*) refers to restructuring, writedown, and M&A activity, respectively, during Q3, and (*post*) refers to this activity during Q4. For *Restate*, due to data constraints, (*pre*) refers to restatement activity during FY−1, and (*post*) refers to this activity during FY, where FY is the fiscal year that is reported on in a firm's initial Section 404 report.

The sample of firms filing an initial Section 404 report consists of firms that (1) provide a Section 404 report in their Form 10-K filings made between November 3, 2004, and March 3, 2006, according to the Audit Analytics Internal Controls database, after eliminating firm-years with erroneous or duplicate CIK codes (these are the firm-identifier codes that Audit Analytics uses) and duplicate or triplicate firm-years (e.g., where a firm disclosed internal control status updates in a subsequent year's Form 10-K), and (2) have all the data needed to estimate our regression models. We remove from the sample 43 firms that report a Section 302 material weakness prior to their initial Section 404 report and four firms for which the absolute value of the test variable *ΔDefaultRisk* (see Table 2) exceeds 400 percent.

$$\begin{aligned}\Delta Spread = {} & \alpha_0 + \beta_0 MW + \beta_1 \Delta DefaultRisk + \beta_2 \Delta Size + \beta_3 \Delta CFO \\ & + \beta_4 \Delta Leverage + \beta_5 \Delta CFVol + \beta_6 \Delta TrsYTM + \beta_7 \Delta Restructure \\ & + \beta_8 \Delta WriteDown + \beta_9 \Delta M\&A + \beta_{10} \Delta Restate + \varepsilon. \qquad (1)\end{aligned}$$

Model (1) employs a difference-in-differences test design that compares each MW and non-MW firm to itself by measuring the change in a firm's credit spread around an initial Section 404 report (*ΔSpread*) as the credit spread following the report minus the credit spread preceding the report. To calculate *ΔSpread*, we first identify all TRACE trades in a firm's debt instruments that occur over days [−45, −3] and days [+3, +45] relative to the initial Section 404 report filing date. For each trade, we measure the credit spread as YTM on the relevant debt issue minus YTM on a matched-by-maturity Treasury security. We then calculate *ΔSpread* for each firm as the mean (for all firm-specific trades) credit spread over days [+3, +45] minus the mean credit spread over days [−45, −3].[14]

By measuring *ΔSpread* as the credit spread following minus the credit spread preceding an initial Section 404 report, we attempt to control for observable firm characteristics that could be used to predict a MW, and thus impact the response of the credit spread to the information in an initial Section 404 report. In particular, *ΔSpread* takes into account the credit spread before a firm files an initial Section 404 report, which includes the effect of these firm characteristics with regards to the debt market's anticipation of how likely it is that a firm will report a MW.

The test design also uses the indicator variable *MW*, set equal to 1 if a firm discloses a MW in an initial Section 404 report, and 0 otherwise, to compare *ΔSpread* of MW firms to *ΔSpread* of non-MW firms. The use of the variable *MW* allows non-MW firms to act as a control for the general increase in credit spreads over the test period that we observe in Table 1. H1 predicts that the estimated coefficient on *MW* is positive.

Model (1) includes controls for factors that previous research finds to influence the cost of debt. We measure the relevant control variables using the change in value (Δ) during the fourth quarter of the fiscal year that is reported on in a firm's initial Section 404 report. Measuring the control variables in this manner helps to alleviate the concern that the reaction of a firm's credit spread to the release of its Form 10-K may be a response to the financial information in the Form 10-K rather than the information in the Section 404 report that is included in the Form 10-K.

---

[14] *ΔSpread* averages (1) multiple trades in a firm's individual debt instruments during a window, and (2) trades in more than one of a firm's debt instruments during a window. In the sample, the mean number of trades in a firm's debt instruments in days [−45, −3] ([+3, +45]) is 210 (204); the median number of trades is 45 (51); and approximately one-third of firms have 20 (27) or fewer trades, while approximately one-third have 100 (110) or more trades. Also, in both windows, the mean (median) number of debt instruments per firm that trade in a window is 2.8 (2.0); and more than 69 percent (96 percent) of firms have one or two (ten or fewer) debt instruments that trade in a window.

To control for an expected positive relationship between the default risk of a firm's debt and its cost of debt, we use the variable $\Delta DefaultRisk$, where *DefaultRisk* is a firm's score from the bankruptcy model of Ohlson (1980).[15] Based on evidence in Begley et al. (1996) and following the procedure in Mansi et al. (2004), we calculate *DefaultRisk* for each firm using the model coefficients provided by Ohlson (1980). We predict that the estimated coefficient on $\Delta DefaultRisk$ is positive.

Previous studies often include controls for firm size, profitability, leverage, and cash flow volatility in models that estimate a firm's cost of debt (e.g., Petersen and Rajan 1994; Sengupta 1998; Kim et al. 2011). For this reason, Model (1) includes the variables $\Delta Size$ (change in the log of total assets), ΔCFO (change in cash flow from operations scaled by total assets), $\Delta Leverage$ (change in long-term debt scaled by total assets), and $\Delta CFVol$ (change in the standard deviation over the last 20 quarters of cash flow from operations scaled by total assets).[16,17] Based on the results of the studies cited above, we predict that the estimated coefficients on $\Delta Size$ and ΔCFO are negative, and the estimated coefficients on $\Delta Leverage$ and $\Delta CFVol$ are positive.

The finance literature documents, theoretically and empirically, that changes in credit spreads are negatively associated with changes in Treasury rates, especially in the short run (e.g., Longstaff and Schwartz 1995; Duffee 1998; Neal et al. 2001). Based on this evidence, Model (1) includes the variable $\Delta TrsYTM$. For each sample firm i, $\Delta TrsYTM_i$ is equal to the mean of all "YTM on a matched-by-maturity Treasury security" used to calculate $Spread_i$ over days [+3, +45] minus the mean of all "YTM on a matched-by-maturity Treasury security" used to calculate $Spread_i$ over days [−45, −3]. Based on the studies cited above, we predict that the estimated coefficient on $\Delta TrsYTM$ is negative.

To control for changes in a firm's financial health, Model (1) includes controls for change in the existence of restructuring, writedown, M&A, and/or restatement activity that would be disclosed in the Form 10-K for the fiscal year of a firm's initial Section 404 report.[18] We do not predict a sign for the estimated coefficients on $\Delta Restructure$ or $\Delta WriteDown$ because restructuring and writedown activity could be bad news (e.g., a firm is having problems) or good news (e.g., a firm is taking action to resolve its problems). Similarly, we do not predict a sign for the estimated coefficient on $\Delta M\&A$ because M&A activity could be bad news (e.g., a firm is overpaying) or good news (e.g., a firm is making a sensible acquisition). Finally, we predict that the estimated coefficient on Δ*Restate* is positive because a restatement indicates weak internal control over financial reporting.

---

[15] The factors in the model are log of assets, working capital scaled by assets, total liabilities scaled by total assets, current liabilities scaled by current assets, net income scaled by assets, operating cash flows scaled by liabilities, change in net income scaled by the sum of the absolute values of the prior two years' net income, and two indicator variables. The indicator variables identify (1) firms with negative earnings in the current and previous year, and (2) firms with negative owner's equity in the current year. If we replace $\Delta DefaultRisk$ with Δ for each of the individual factors included in the bankruptcy model of Ohlson (1980), the only notable change in our results is that the estimated coefficient on *MW* in Model (1) is now positive at $p = 0.153$ (rather than $p = 0.099$).

[16] If we use a measure of ΔCFO that attempts to control for seasonality (i.e., ΔCFO as currently measured minus ΔCFO for the fiscal year preceding the fiscal year that is reported on in a firm's initial Section 404 report), it is significantly correlated with ΔCFO ($p = 0.001$) and there is no notable change in our results.

[17] We do not use change in earnings to measure change in profitability because, as a result of earnings announcements made prior to the release of the Form 10-K, it is likely that credit spreads would have already reacted to the information in change in earnings during the days [−45, −3] window. If we replace ΔCFO with "actual annual EPS reported by I/B/E/S minus the most recent (but no more recent than the end of the third quarter) mean I/B/E/S forecast for annual EPS," there is no notable change in our results.

[18] $\Delta Restructure$, $\Delta WriteDown$, $\Delta M\&A$, and $\Delta Restate$ are coded as −1, 0, or 1, as described in Table 2.

Test Results

As reported in Table 2, when we estimate Model (1), the coefficient on *MW* is positive ($p = 0.099$).[19] This result provides marginal support for H1 that the disclosure of a MW in an initial Section 404 report is associated with an increase in a firm's cost of debt.[20] However, this initial result is an average result. It does not provide evidence as to whether the effect of disclosing a MW on the cost of debt holds equally for firms that are monitored and firms that are not monitored by credit rating agencies and/or banks. Therefore, we next explore the impact of monitoring.

Tests of H2a and H2b: Monitoring Effects

Test Design

To examine the monitoring hypotheses, which predict that the evidence supporting H1 is stronger for firms that are not monitored by credit rating agencies and/or banks compared to firms that are monitored, we estimate Model (2):

$$\Delta Spread = \alpha_0 + \beta_0 MW + \beta_1 Monitored + \beta_2 MW * Monitored + \beta_n Model\ 1\ Controls + \varepsilon. \quad (2)$$

Model (2) employs a difference in difference-in-differences test design. The innovation in Model (2) is that the coefficient β_2 tests the difference between (1) the difference in Δ*Spread* between MW and non-MW firms for monitored firms, and (2) the difference in Δ*Spread* between MW and non-MW firms for unmonitored firms. This innovation allows us to test whether the evidence supporting H1 is stronger for unmonitored firms. The monitoring hypotheses predict that the estimated coefficient β_2 is negative.

Test Results for Monitoring by Credit Rating Agencies

To test the effect of monitoring by credit rating agencies, we measure *Monitored* using an indicator variable *Rated*, set equal to 1 if a firm's debt is rated by S&P and/or Moody's, and 0 otherwise. To determine if S&P rates a firm's debt, we use the S&P credit rating variable available in Compustat. Further, to determine if Moody's rates a firm's debt, we consult the Mergent Bond Record as of January 2005 (Mergent Inc. 2005), which lists debt issues that Moody's rates.

In our sample of 577 firms, there are 489 firms with rated debt and 88 firms with unrated debt.[21] Among the 489 rated firms, there are 456 non-MW and 33 MW firms. Among the 88 unrated firms, there are 75 non-MW and 13 MW firms. As reported in Table 3, Panel A, when we estimate Model (2) using *Rated*, the coefficient β_2 is negative ($p = 0.039$), supporting H2a that the

---

[19] We report p-values of 0.001 or better as $p = 0.001$. Reported p-values are based on one-tailed tests. In our tests reported in Tables 3 through 6, the significant results for the control variables with a predicted sign are that the estimated coefficient on Δ*CFO* is negative ($p = 0.001$), the estimated coefficient on Δ*CFVol* is positive (between $p = 0.003$ and $p = 0.062$), and the estimated coefficient on Δ*TrsYTM* is negative (between $p = 0.001$ and $p = 0.011$).

[20] Consistent with this finding, in univariate tests, mean Δ*Spread* for MW firms (non-MW firms) is 0.687 (0.368), and the difference in mean Δ*Spread* between the two groups of firms is positive and statistically significant ($p = 0.047$).

[21] Choi and Richardson (2010), in Panel A of Table 2, report that among their 1,566 sample firms with publicly traded bonds, 266 firms (16.99 percent) are unrated, which is close to our finding that 15.25 percent of our sample firms are unrated.

TABLE 2

Analysis of the Change in a Firm's Credit Spread following an Initial Section 404 Report

ΔSpread =			Coefficient	p-value[a]
Constant			0.717	0.001
MW	β_0	+	0.242	0.099
ΔDefaultRisk	β_1	+	−0.051	0.782
ΔSize	β_2	−	0.369	0.787
ΔCFO	β_3	−	−6.025	0.001
ΔLeverage	β_4	+	0.338	0.365
ΔCFVol	β_5	+	19.192	0.005
ΔTrsYTM	β_6	−	−1.084	0.001
ΔRestructure	β_7		0.234	0.107
ΔWriteDown	β_8		−0.343	0.033
ΔM&A	β_9		−0.053	0.457
ΔRestate	β_{10}	+	−0.143	0.899
Adjusted R^2			0.080	

---

[a] p-values are one-tailed, and p = 0.001 indicates that p is less than or equal to 0.001. p-values exceeding 0.500 indicate that the sign of an estimated coefficient is not in the predicted direction.

The variables (followed by predicted signs) in the OLS regression model reported in this table are as follows.

Dependent variable: The change in a firm's credit spread around an initial Section 404 report (*ΔSpread*), computed as follows. We measure the credit spread for individual trades in a firm's debt instruments as YTM on the relevant debt issue minus YTM on a matched-by-maturity Treasury security. We calculate the mean pre- (post-) report credit spread for each firm (i.e., *Spread*) as the mean credit spread for all TRACE-reported trades occurring over days [−45, −3] ([+3, +45]) relative to the initial Section 404 report filing date. We calculate the change in credit spread for each firm as its mean post-report credit spread minus its mean pre-report credit spread.

RHS variables:
MW = 1 if a firm discloses a MW in an initial Section 404 report, and 0 otherwise;
ΔDefaultRisk = change in the score from the bankruptcy model of Ohlson (1980);
ΔSize = change in the log of total assets;
ΔCFO = change in cash flow from operations scaled by total assets;
ΔLeverage = change in long-term debt scaled by total assets;
ΔCFVol = change in standard deviation over the last 20 quarters of cash flow from operations scaled by total assets, where change in these "Δ variables" is change in value during the fourth quarter of the fiscal year that is reported on in a firm's initial Section 404 report (Q4);
For each sample firm *i*, $\Delta TrsYTM_i$ = the mean of all "YTM on a matched-by-maturity Treasury security" used to calculate $Spread_i$ over days [+3, +45] minus the mean of all "YTM on a matched-by-maturity Treasury security" used to calculate $Spread_i$ over days [−45, −3];
ΔRestructure = 1 if restructuring activity occurs in Q4 but not in the preceding quarter, Q3 (equal to −1 if restructuring activity occurs in Q3 but not in Q4; equal to 0 if restructuring activity occurs in both Q3 and Q4 or if no restructuring activity occurs in either Q3 or Q4); and
ΔWriteDown, *ΔM&A*, and *ΔRestate* are measured in a manner similar to *ΔRestructure* to reflect changes in writedown, M&A, and restatement activity, respectively. Due to data constraints, *ΔRestate* is measured for the fiscal year (rather than Q4) that is reported on in a firm's initial Section 404 report.
The construction of the sample of 577 (531 non-MW and 46 MW) firms is described in Table 1.

increased cost of debt associated with MW disclosure is more pronounced for firms with unrated debt than for firms with rated debt.

Next, we assess the impact of reporting a MW on *ΔSpread* for (1) only firms with rated debt, and (2) only firms with unrated debt. In terms of Model (2), the relevant tests are based on (1) the sum of the estimated coefficients $\beta_0 + \beta_2$ for firms with rated debt, and (2) the estimated coefficient β_0 for firms with unrated debt. As reported in Table 3, Panel A, the sum of the estimated

TABLE 3

Analysis of the Change in a Firm's Credit Spread Following an Initial Section 404 Report with Independent Controls for Monitoring

Panel A: Control for Monitoring by Credit Rating Agencies

ΔSpread =			**Coefficient**	**p-value**[a]
Constant			0.565	0.001
MW	β_0	+	0.798	0.014
Rated	β_1		−0.193	0.114
MW * *Rated*	β_2	−	−0.742	0.039
ΔDefaultRisk	β_3	+	−0.060	0.818
ΔSize	β_4	−	0.289	0.731
ΔCFO	β_5	−	−6.307	0.001
ΔLeverage	β_6	+	0.410	0.340
ΔCFVol	β_7	+	20.879	0.003
ΔTrsYTM	β_8	−	−1.098	0.001
ΔRestructure	β_9		0.228	0.112
ΔWriteDown	β_{10}		−0.339	0.034
ΔM&A	β_{11}		−0.075	0.439
ΔRestate	β_{12}	+	−0.155	0.917
Adjusted R^2			0.083	
Sum of coefficients				
MW + *MW* * *Rated* ($\beta_0 + \beta_2$)		+	0.056	0.400

Panel B: Control for Monitoring by Banks

ΔSpread =			**Coefficient**	**p-value**[a]
Constant			0.796	0.001
MW	β_0	+	1.434	0.001
Bank	β_1		−0.119	0.235
MW * *Bank*	β_2	−	−1.744	0.001
ΔDefaultRisk	β_3	+	−0.038	0.722
ΔSize	β_4	−	0.190	0.662
ΔCFO	β_5	−	−5.079	0.001
ΔLeverage	β_6	+	−0.211	0.586
ΔCFVol	β_7	+	15.915	0.016
ΔTrsYTM	β_8	−	−1.111	0.001
ΔRestructure	β_9		0.270	0.072
ΔWriteDown	β_{10}		−0.283	0.062
ΔM&A	β_{11}		−0.060	0.451
ΔRestate	β_{12}	+	−0.116	0.854
Adjusted R^2			0.114	
Sum of coefficients				
MW + *MW* * *Bank* ($\beta_0 + \beta_2$)		+	−0.310	0.922

---

[a] p-values are one-tailed, and $p = 0.001$ indicates that p is less than or equal to 0.001. p-values exceeding 0.500 indicate that the sign of an estimated coefficient is not in the predicted direction.

The variables (followed by predicted signs) in the OLS regression models reported in this table are as follows:

Dependent Variable: The change in a firm's credit spread around an initial Section 404 report (*ΔSpread*) is defined in Table 2.

(continued on next page)

TABLE 3 (continued)

RHS variables:

Panel A: *Rated* = 1 if a firm's debt is rated by S&P and/or Moody's (489 firms: 456 are non-MW firms and 33 are MW firms), and 0 otherwise (88 firms: 75 are non-MW firms and 13 are MW firms).

Panel B: *Bank* = 1 if a firm has a new or revised bank loan agreement in any year from 1988 to 2005 (500 firms: 468 are non-MW firms and 32 are MW firms), and 0 otherwise (77 firms: 63 are non-MW firms and 14 are MW firms).

Panels A and B: *MW*, *ΔDefaultRisk*, *ΔSize*, *ΔCFO*, *ΔLeverage*, *ΔCFVol*, *ΔTrsYTM*, *ΔRestructure*, *ΔWriteDown*, *ΔM&A*, and *ΔRestate* are defined in Table 2.

The construction of the sample of 577 (531 non-MW and 46 MW) firms is described in Table 1.

coefficients $\beta_0 + \beta_2$ is positive but not statistically significant ($p = 0.400$), and the estimated coefficient β_0 is positive and statistically significant ($p = 0.014$). These results provide further support for H2a.[22] Initially, these findings suggest that monitoring by credit rating agencies is effective at capturing much of the information provided to the debt market in an initial Section 404 report.

Test Results for Monitoring by Banks

We now turn to tests of the bank-monitoring hypothesis. To test the effect of monitoring by banks, we measure *Monitored* using an indicator variable *Bank*. To code *Bank*, we first consult the DealScan database to identify sample firms that have a new or revised bank loan agreement in any year from 1988 (the earliest DealScan record for a sample firm) through 2005. We code these firms *Bank* = 1 (i.e., subject to bank monitoring).[23] We then code the remaining sample firms *Bank* = 0.

In our sample of 577 firms, there are 500 firms that are coded *Bank* = 1 (i.e., subject to bank monitoring) and 77 firms that are coded *Bank* = 0. Among the 500 bank-monitored firms, there are 468 non-MW and 32 MW firms. Among the 77 non-bank-monitored firms, there are 63 non-MW and 14 MW firms. Referring to Table 3, Panel B, in the estimate of Model (2) using *Bank*, the coefficient β_2 is significantly negative ($p = 0.001$). Thus, the data provide support for H2b that the increased cost of debt associated with MW disclosure is more pronounced for firms that are not subject to bank monitoring than for firms that are subject to bank monitoring.

Next, we assess the impact of reporting a MW on *ΔSpread* for (1) only bank-monitored firms, and (2) only firms not subject to bank monitoring. As reported in Table 3, Panel B, in this analysis the sum of the estimated coefficients $\beta_0 + \beta_2$ is negative and the estimated coefficient β_0 is positive and statistically significant ($p = 0.001$). These results provide further support for H2b.[24]

---

[22] Univariate tests highlight the importance of monitoring by credit rating agencies. For rated firms, mean *ΔSpread* for MW firms (non-MW firms) is 0.453 (0.351), and the difference in mean *ΔSpread* between the two groups of firms is positive but not significant at conventional levels ($p = 0.305$). For unrated firms, mean *ΔSpread* for MW firms (non-MW firms) is 1.281 (0.476), and the difference in mean *ΔSpread* between the two groups of firms is marginally significant ($p = 0.067$).

[23] The DealScan database is compiled using information from SEC filings and self-reported information from participating banks. Some loan packages or deals reported by DealScan can have several facilities for the same borrower. We do not break down such loan packages or deals into individual facilities because we are using DealScan to identify the existence of a new or revised bank loan agreement, which could be a single facility, a loan package, or a deal with several individual facilities. Also, it is not necessarily the case that a new or revised bank loan agreement entered into prior to the initial Section 404 report filing date will still be in effect when the report is filed; we analyze this issue in detail in Section VI.

[24] Univariate tests highlight the importance of bank monitoring. For bank-monitored firms, mean *ΔSpread* for MW firms (non-MW firms) is 0.079 (0.326), and the difference in mean *ΔSpread* between the two groups of firms is negative. For firms not subject to bank monitoring, mean *ΔSpread* for MW firms (non-MW firms) is 2.076 (0.681), and the difference in mean *ΔSpread* between the two groups of firms is positive and statistically significant ($p = 0.014$).

The preceding findings support both the credit rating agency and the bank-monitoring hypotheses. However, a test is still required to assess whether each of the monitoring effects documented so far has an incremental effect on the relationship between the change in a firm's credit spread and the disclosure of a MW. Therefore, we next test the joint effects of monitoring by credit rating agencies and banks.

Test Results for the Joint Effects of Monitoring

In this section, we test H2a and H2b jointly, using Model (3):

$$\begin{aligned}\Delta Spread = \alpha_0 &+ \beta_0 MW + \beta_1 Rated + \beta_2 MW * Rated + \beta_3 Bank + \beta_4 MW * Bank \\ &+ \beta_5 Rated * Bank + \beta_6 MW * Rated * Bank + \beta_n Model\,1\ Controls + \varepsilon. \qquad (3)\end{aligned}$$

Model (3) allows us to test the difference between (1) the difference in $\Delta Spread$ between MW and non-MW firms that have rated debt and are subject to bank monitoring (i.e., the "most monitored" firms), and (2) the difference in $\Delta Spread$ between MW and non-MW firms that have unrated debt and are not subject to bank monitoring (i.e., the "least monitored" firms). To compare the "most monitored" and the "least monitored" firms, we test the sum of the estimated coefficients β_2, β_4, and β_6. The monitoring hypotheses predict that the sum of these coefficients is negative. As reported in Table 4, this prediction is supported ($p = 0.002$).[25]

Model (3) also allows us to test whether the monitoring effect of credit rating agencies and that of banks each has its own incremental effect on the relationship between the change in a firm's credit spread and the disclosure of a MW, as opposed to one effect subsuming the other.[26] In our sample, there are 454 firms monitored by both credit rating agencies and banks, 35 firms monitored only by credit rating agencies, 46 firms monitored only by banks, and 42 firms monitored by neither.[27] The relevant tests, predictions, and results (reported in Table 4) comparing these subgroups are as follows:

- For bank-monitored firms, is there a significant difference between (1) the difference in $\Delta Spread$ between MW and non-MW firms for rated firms, and (2) the difference in $\Delta Spread$ between MW and non-MW firms for unrated firms? Relevant prediction (H2a): the sum of the estimated coefficients β_2 and β_6 is negative. Result: the data fail to support this prediction ($p = 0.261$).
- For firms not subject to bank monitoring, is there a significant difference between (1) the difference in $\Delta Spread$ between MW and non-MW firms for rated firms, and (2) the difference in $\Delta Spread$ between MW and non-MW firms for unrated firms? Relevant prediction (H2a): the estimated coefficient β_2 is negative. Result: the data fail to support this prediction (i.e., the estimated coefficient β_2 is positive).

P 32,763

---

[25] Univariate tests highlight the importance of monitoring. For firms that have rated debt and are subject to bank monitoring, mean $\Delta Spread$ for MW firms (non-MW firms) is 0.038 (0.335), and the difference in mean $\Delta Spread$ between the two groups of firms is negative. For firms that have unrated debt and are not subject to bank monitoring, mean $\Delta Spread$ for MW firms (non-MW firms) is 1.893 (0.774), and the difference in mean $\Delta Spread$ between the two groups of firms is positive and marginally significant ($p = 0.084$).

[26] For example, if a firm is both rated and subject to bank monitoring, it is possible that the contribution of this firm to the significant "monitoring matters" results could be caused by the fact that the firm is rated, by the fact that the firm is bank-monitored, or by both. In other words, when we measure *Monitored* in Model (2) using the indicator variable *Rated* (*Bank*), *Bank* (*Rated*) may be a correlated omitted variable. Model (3) attempts to get at this type of issue.

[27] Among the 454 firms monitored by both credit rating agencies and banks, there are 427 non-MW and 27 MW firms. Among the 35 firms monitored only by credit rating agencies, there are 29 non-MW and six MW firms. Among the 46 firms monitored only by banks, there are 41 non-MW and five MW firms. Among the 42 firms monitored by neither, there are 34 non-MW and eight MW firms. In Section VII, we discuss the fact that some of the MW sample cells used in these joint tests are quite small.

TABLE 4

Analysis of the Change in a Firm's Credit Spread Following an Initial Section 404 Report with Joint Controls for Monitoring

ΔSpread =			Coefficient	p-value[a]
Constant			0.781	0.001
MW	β_0		1.161	0.007
Rated	β_1		0.038	0.451
*MW * Rated*	β_2	–	0.612	0.806
Bank	β_3		−0.349	0.107
*MW * Bank*	β_4	–	−1.132	0.059
*Rated * Bank*	β_5		0.249	0.244
*MW * Rated * Bank*	β_6		−0.998	0.143
ΔDefaultRisk	β_7	+	−0.053	0.790
ΔSize	β_8	–	0.229	0.690
ΔCFO	β_9	–	−5.342	0.001
ΔLeverage	β_{10}	+	0.000	0.500
ΔCFVol	β_{11}	+	16.509	0.015
ΔTrsYTM	β_{12}	–	−1.143	0.001
ΔRestructure	β_{13}		0.281	0.065
ΔWriteDown	β_{14}		−0.274	0.068
ΔM&A	β_{15}		−0.095	0.422
ΔRestate	β_{16}	+	−0.112	0.845
Adjusted R^2			0.113	
Sums of coefficients				
*MW * Rated + MW * Bank + MW * Rated * Bank* ($\beta_2 + \beta_4 + \beta_6$)		–	−1.518	0.002
*MW * Rated + MW * Rated * Bank* ($\beta_2 + \beta_6$)		–	−0.386	0.261
*MW * Bank + MW * Rated * Bank* ($\beta_4 + \beta_6$)		–	−2.130	0.001

---

[a] p-values are one-tailed, and p = 0.001 indicates that p is less than or equal to 0.001. p-values exceeding 0.500 indicate that the sign of an estimated coefficient is not in the predicted direction.

The variables (followed by predicted signs) in the OLS regression model reported in this table are as follows:

Dependent variable: The change in a firm's credit spread around an initial Section 404 report (*ΔSpread*) is defined in Table 2.

RHS variables: *MW*, *ΔDefaultRisk*, *ΔSize*, *ΔCFO*, *ΔLeverage*, *ΔCFVol*, *ΔTrsYTM*, *ΔRestructure*, *ΔWriteDown*, *ΔM&A*, and *ΔRestate* are defined in Table 2. *Rated* and *Bank* are defined in Table 3.

The construction of the sample of 577 (531 non-MW and 46 MW) firms is described in Table 1.

- For rated firms, is there a significant difference between (1) the difference in *ΔSpread* between MW and non-MW firms for bank-monitored firms, and (2) the difference in *ΔSpread* between MW and non-MW firms for firms not subject to bank monitoring? Relevant prediction (H2b): the sum of the estimated coefficients β_4 and β_6 is negative. Result: the data provide strong support for this prediction (p = 0.001).
- For unrated firms, is there a significant difference between (1) the difference in *ΔSpread* between MW and non-MW firms for bank-monitored firms, and (2) the difference in *ΔSpread* between MW and non-MW firms for firms not subject to bank monitoring? Relevant prediction (H2b): the estimated coefficient β_4 is negative. Result: the data provide marginal support for this prediction (p = 0.059).

The preceding tests suggest that bank monitoring subsumes monitoring by credit rating agencies with regards to the post-report response of credit spread to the disclosure of a MW in an initial Section 404 report. Of primary interest are the third and fourth results, which suggest that bank monitoring reduces the impact of reporting a MW in an initial Section 404 report regardless of credit rating status. These results suggest that the effect of bank monitoring may be the primary driver of our more general "monitoring matters" results.

Test of H3: Delinquent Filers

We now test whether there is a difference in the change in credit spread between firms that are delinquent in filing an initial Section 404 report and timely filers. We classify a firm as a delinquent filer if it does not file an initial Section 404 report within 90 days of its fiscal year-end.[28] Conversely, a timely filer is defined as a firm that files an initial Section 404 report within 90 days of its fiscal year-end.

For each delinquent and timely filer, we measure the change in credit spread (Δ*Spread*(*adjusted*)) around this 90-day cutoff as the mean credit spread for all TRACE-reported trades in a firm's debt issues over days [+3, +45] minus the mean credit spread for all TRACE-reported trades over days [−45, −3] relative to the 90-day cutoff. We then use the 544 sample firms (533 timely filers and 11 delinquent filers) with complete data availability, as of the 90-day cutoff, to estimate a regression of Δ*Spread*(*adjusted*) on *Delinquent* and the control variables included in our previous models. *Delinquent* is equal to 1 if a firm is a delinquent filer, and 0 otherwise. As reported in Table 5, the estimated coefficient on *Delinquent* is significantly positive ($p = 0.001$).[29] Consistent with Ogneva et al.'s (2007) finding for the cost of equity, the result for *Delinquent* indicates that delinquent filers experience a significant increase in credit spread relative to timely filers.

V. COMPARISON TO THE COST OF EQUITY RESULTS IN ASHBAUGH-SKAIFE ET AL. (2009)

In this section, we compare our findings to those of Ashbaugh-Skaife et al. (2009), which focus on the cost of equity rather than the cost of debt. We begin by partitioning the data by the *ex ante* probability of reporting a MW, which allows for a more direct comparison between the two studies.[30] To partition the data, we first estimate the logistic probability model in Ashbaugh-Skaife et al. (2007), using all the firms we identified as releasing an initial Section 404 report between November 3, 2004, and March 3, 2006, that have the data necessary to estimate the model.[31] Next, we determine the overlap between the firms in the upper and lower halves of the resulting *ex ante* probability of reporting a MW distribution and the 577 firms in the sample used to estimate our earlier models. These procedures result in a sample of 506 firms: 24 MW (134 non-MW) firms in

---

[28] The SEC filing requirement in effect for Form 10-Ks during our sample period is 75 days for accelerated filers, with an extra 15 days allowed to file without penalty. SEC Release No. 34-50754 (SEC 2004) grants an additional 45 days beyond the 75-day filing requirement for small accelerated filers to file a Section 404 report. We choose the 90-day cutoff because we view failing to meet this deadline as an appropriate indicator of pervasive internal control over financial reporting problems.

[29] Consistent with this result, in univariate tests, mean Δ*Spread* for delinquent firms (timely filers) is 2.097 (0.532), and the difference in mean Δ*Spread* between the two groups of firms is positive and statistically significant ($p = 0.001$).

[30] We choose Ashbaugh-Skaife et al.'s (2009) results as our benchmark because their test design is similar to ours. Ashbaugh-Skaife et al.'s (2009) results do not provide a strict "apples to apples" comparison, primarily because they examine the disclosure of a Section 302 MW or a Section 404 MW.

[31] The determinants in the model are number of business segments, decile rank of sales growth, inventory to assets ratio, proportion of three preceding years with negative earnings, decile rank of Altman (1968) z-score, percent of shares held by institutional investors divided by number of institutional owners, market value of equity, and indicator variables for foreign sales, M&A activity, restructuring, auditor resignation, earnings restatement, membership in a litigious industry, and use of a large auditor.

TABLE 5

Analysis of the Change in a Firm's Credit Spread: Timely versus Delinquent Filers

$\Delta Spread(adj)$ =		Coefficient	p-value[a]
Constant		0.801	0.001
$Delinquent$	+	1.485	0.001
$\Delta DefaultRisk$	+	−0.031	0.653
$\Delta Size$	−	0.399	0.759
ΔCFO	−	−5.946	0.001
$\Delta Leverage$	+	0.593	0.318
$\Delta CFVol$	+	13.747	0.062
$\Delta TrsYTM$	−	−0.518	0.011
$\Delta Restructure$		0.005	0.490
$\Delta WriteDown$		−0.283	0.096
$\Delta M\&A$		−0.164	0.385
$\Delta Restate$	+	−0.009	0.527
Adjusted R^2		0.046	

---

[a] p-values are one-tailed and $p = 0.001$ indicates that p is less than or equal to 0.001. p-values exceeding 0.500 indicate that the sign of the estimated coefficient is not in the predicted direction.

The variables (followed by predicted signs) in the OLS regression model reported in this table are as follows:

Dependent variable: The change in a firm's credit spread ($\Delta Spread(adj)$), computed as follows. We measure the credit spread for individual trades in a firm's debt instruments as YTM on the relevant debt issue minus YTM on a matched-by-maturity Treasury security. We calculate the mean pre- (post-) cutoff date credit spread for each firm as the mean credit spread for all TRACE-reported trades occurring over days [−45, −3] ([+3, +45]) relative to the cutoff of 90 days after a firm's fiscal year-end. We calculate the change in credit spread for each firm as its mean post-cutoff credit spread minus its mean pre-cutoff credit spread.

RHS variables: $Delinquent = 1$ if a firm does not file an initial Section 404 report within 90 days of its fiscal year-end, and 0 otherwise. $\Delta DefaultRisk$, $\Delta Size$, ΔCFO, $\Delta Leverage$, $\Delta CFVol$, $\Delta TrsYTM$, $\Delta Restructure$, $\Delta WriteDown$, $\Delta M\&A$, and $\Delta Restate$ are defined in Table 2.

Among the 577 sample firms described in Table 1 that have complete data availability as of the 90-day cutoff date for the Table 5 tests, there are 533 firms that file a timely report, and 11 firms that are delinquent filers.

the upper half of the *ex ante* probability of reporting a MW distribution, and 16 MW (332 non-MW) firms in the lower half. The 158 firms in the upper half of the distribution are coded $HighProb = 1$ and the 348 firms in the lower half are coded $HighProb = 0$.

For these 506 firms, we estimate a regression of $\Delta Spread$ on MW, $HighProb$, $MW * HighProb$, and the control variables included in our previous models. As reported in Table 6, the estimated coefficient on $MW * HighProb$ is significantly negative ($p = 0.001$). This result is consistent with the notion that firms with a low *ex ante* probability of reporting a MW have a larger increase in credit spread in response to reporting a MW than firms with a high *ex ante* probability because the debt market is more surprised when a low *ex ante* probability firm reveals a MW.

Comparing our test results to those in Ashbaugh-Skaife et al. (2009), similar to their finding that a firm's cost of equity increases if it discloses a MW, we find that the disclosure of a MW is associated with an increase in a firm's cost of debt. Also, both studies find evidence consistent with the notion that firms with a low *ex ante* probability of reporting a MW have a larger increase in their cost of capital in response to reporting a MW than high *ex ante* probability firms.

When we partition our sample firms based on whether a firm is monitored by credit rating agencies and/or banks, we observe that the increased cost of debt associated with MW disclosure is more pronounced for firms that are not monitored than for firms that are monitored. Ashbaugh-Skaife et al. (2009) do not examine the impact of monitoring on the cost of equity. An interesting issue for future research to examine is whether there are alternate monitoring

TABLE 6

Analysis of the Change in a Firm's Credit Spread Following an Initial Section 404 Report with Control for Probability of Reporting a MW

Δ*Spread* =		Coefficient	p-value[a]
Constant		0.630	0.001
MW		1.358	0.001
HighProb		0.299	0.017
MW * *HighProb*	−	−1.756	0.001
Δ*DefaultRisk*	+	0.001	0.493
Δ*Size*	−	0.275	0.714
Δ*CFO*	−	−6.180	0.001
Δ*Leverage*	+	0.513	0.315
Δ*CFVol*	+	22.789	0.003
Δ*TrsYTM*	−	−1.035	0.001
Δ*Restructure*		0.360	0.038
Δ*WriteDown*		−0.498	0.007
Δ*M&A*		0.017	0.486
Δ*Restate*	+	−0.118	0.829
Adjusted R^2		0.119	

---

[a] p-values are one-tailed and p = 0.001 indicates that p is less than or equal to 0.001. p-values exceeding 0.500 indicate that the sign of an estimated coefficient is not in the predicted direction.

The variables (followed by predicted signs) in the OLS regression model reported in this table are as follows:

Dependent variable: The change in a firm's credit spread around an initial Section 404 report (Δ*Spread*) is defined in Table 2.

RHS variables: *MW*, Δ*DefaultRisk*, Δ*Size*, Δ*CFO*, Δ*Leverage*, Δ*CFVol*, Δ*TrsYTM*, Δ*Restructure*, Δ*WriteDown*, Δ*M&A*, and Δ*Restate* are defined in Table 2.

Using all the firms we identified as releasing an initial Section 404 report between November 3, 2004, and March 3, 2006, that have the necessary data, we use the logistic probability model in Ashbaugh-Skaife et al. (2007) to estimate the *ex ante* probability of reporting a MW. Next, we determine the overlap between the firms in the upper and lower halves of the resulting *ex ante* probability of reporting a MW distribution and the 577 firms in the sample described in Table 1, resulting in 24 MW (134 non-MW) firms in the upper half of the distribution and 16 MW (332 non-MW) firms in the lower half. These 506 firms are used to estimate the model in Table 6. For these 506 firms, the 158 firms in the upper half of the *ex ante* probability of reporting a MW distribution are coded *HighProb* = 1, and the 348 firms in the lower half are coded *HighProb* = 0.

mechanisms in the equity market (e.g., analyst following) that impact the relationship between the change in a firm's cost of equity and the disclosure of a MW. Another issue future research might consider is a test of the differential effect of various monitoring agents (e.g., banks, credit rating agencies, audit committees, an independent press) on (1) the change in a firm's cost of equity and the disclosure of a MW, and (2) the change in a firm's cost of debt and the disclosure of a MW.

VI. ADDITIONAL ANALYSES

Tests that Classify MW Firms by Severity of MW

Moody's Investors Service, Inc. (2004) classifies a Section 404 MW as either Category A or Category B, as follows:

- A Category A MW relates to controls over specific accounts or transaction-level processes. In general, Moody's does not bring firms with a Category A MW to their rating committee to determine if a credit rating downgrade is necessary.

- A Category B MW relates to company-level control factors such as the control environment. In general, Moody's brings firms with a Category B MW to their rating committee to determine if a credit rating downgrade is necessary.

A Special Comment of Moody's Investors Service, Inc. (2005b) identifies firms with Category A or Category B MW. The Special Comment discusses 18 of our 46 sample MW firms. Among these 18 firms, 11 are classified as Category A and seven are classified as Category B. To determine if MW classification (i.e:, as Category A or Category B) impacts the post-initial Section 404 report change in credit spread for MW firms, we use the 18 MW firms just described to estimate a regression of *ΔSpread* on *ABIndicator* (equal to 1 for Category B firms, and 0 for Category A firms) and the control variables included in our previous models. In this estimate (not tabled), the coefficient on *ABIndicator* is positive, as expected, but not statistically significant (p = 0.274).

Additional Analysis of DealScan Data

It is not necessarily the case that a new or revised bank loan agreement entered into prior to the initial Section 404 report filing date will still be in effect when the report is filed. Thus, it is possible that the variable *Bank*, coded as 1 if a firm has a new or revised bank loan agreement in any year from 1988 through 2005, and 0 otherwise, has measurement error. In particular, some firms, coded *Bank* = 1, may no longer be subject to bank monitoring because all previous bank loan agreements identified on DealScan have expired as of the initial Section 404 report filing date.[32] In this section, we analyze the DealScan data to assess the sensitivity of our results to this issue.

When we conduct a detailed analysis of the DealScan data for our firms, 86 percent of the firms coded as subject to bank monitoring have at least one new or revised bank loan agreement in the 2003 to 2005 window. Also, among the 14 percent of firms whose most recent new or revised bank loan agreement occurs before 2003, for all but ten of these firms, the most recent agreement is either a long-term revolving line of credit or a term loan.[33] Thus, for all but ten sample firms, the most recent new or revised bank loan agreement either (1) occurs in a relatively short window preceding the initial Section 404 report filing date, or (2) involves a loan that is presumably of a long-term nature. We interpret these observations as suggesting that it is reasonable to code *Bank* using our long coding window, especially given the miscoding issue involved using a short coding window (see footnote 32).

Tests of Change in Credit Rating

Several previous studies examine the relationship between credit rating and MW status. Using cross-sectional tests, Elbannan (2009) finds that firms that disclose a Section 302 MW or a Section 404 MW have relatively low credit ratings. Using changes tests, Crabtree et al. (2009) find that the disclosure of a MW in an initial Section 404 report is associated with a rating downgrade. In this section, we test whether it is the case for our sample firms that the disclosure of a MW in an initial Section 404 report is associated with a rating downgrade. Our tests regress *ΔRating* on the variables included in our previous models, where *ΔRating* is the change in a firm's credit rating after a Section 404 report is filed.[34]

---

[32] We note that the choice of a long coding window is based on a trade-off. If a short window is used, firms that enter into a still-active bank loan agreement prior to the short window will be miscoded *Bank* = 0 (as opposed to the miscoded *Bank* = 1 firms that result from using a long window).

[33] If we remove the ten problem firms from the sample and estimate our models that include the variable *Bank*, there is no notable change in our results.

In this analysis, there are 480 (450 non-MW and 30 MW) firms from the original sample that have the necessary ratings data. We are unable to estimate models that include the variable *Rated* because all firms in these tests are rated. Table 7 summarizes the distribution of MW and non-MW observations across various credit rating classes in both the pre-Section 404 report period and the post-Section 404 report period. Consistent with the claim of credit rating agencies that their ratings reflect internal control deficiencies before they are eventually disclosed in a Section 404 report (e.g., Standard & Poor's [S&P] 2004; Moody's Investors Service, Inc. 2005b), MW firms tend to fall within the lower credit rating classes in the pre-Section 404 report period. Also, in the pre-Section 404 report period, the mean credit rating for MW firms is significantly worse than that of non-MW firms ($p = 0.001$; untabulated).

In our model estimates using *ΔRating* (untabulated), the coefficient on *MW* in revised Model (1) is not statistically significant ($p = 0.118$). In revised Model (2), the coefficient on *MW * Bank*, the coefficient on *MW*, and the sum of the coefficients on *MW* and *MW * Bank* each fail to reach statistical significance (most significant $p = 0.214$). These results suggest that, consistent with the claim of credit rating agencies that their ratings reflect internal control deficiencies before they are eventually disclosed in a Section 404 report, credit ratings do not appear to be very sensitive to the disclosure of a MW in an initial Section 404 report.

Comparison to the Results in Kim et al. (2011) and Costello and Wittenberg-Moerman (2011)

Concurrent studies by Kim et al. (2011) and Costello and Wittenberg-Moerman (2011) use bank loan credit spreads to examine some of the same issues investigated in our study. Both Kim et al. (2011) and Costello and Wittenberg-Moerman (2001) find that there is an increase in a firm's cost of debt for new bank loans initiated after reporting a MW. Based on these findings, we view Kim et al. (2011), Costello and Wittenberg-Moerman (2011), and our study as complementary in that they collectively provide evidence that Sarbanes-Oxley reporting requirements can impact a firm's cost of debt in both the private debt market and the public debt market.

We note here that there are many differences in test design among the three studies described in this section. To provide three examples, we observe the following:

- In difference-in-differences tests, our study focuses on changes in the cost of debt for *existing* public debt issues, while Kim et al. (2011) and Costello and Wittenberg-Moerman (2011) focus on changes in the cost of debt for *new* bank loans initiated after reporting a MW.
- In difference-in-differences tests, the test window in our study is 90 days, the test window in Kim et al. (2011) is up to four years, and the test window in Costello and Wittenberg-Moerman (2011) exceeds three years.
- In our study and Kim et al. (2011), the major focus is on Section 404 disclosures, while in Costello and Wittenberg-Moerman (2011), the major focus is on Section 302 disclosures.

Despite differences in test design, all three studies described in this section find evidence that a firm's cost of debt increases after it discloses a MW, suggesting that one benefit of Sarbanes-Oxley

---

[34] Credit rating data are from the Compustat Ratings File, which reports monthly long-term issuer credit ratings for firms in the S&P 1500. We measure *ΔRating* as *Rating* at the end of month $m + 1$ minus *Rating* at the beginning of month $m - 1$, where m is the initial Section 404 report filing month. *Rating* in a given month is coded as in the Compustat Ratings File (e.g., AAA = 2, AA+ = 4, AA = 5, AA− = 6, A+ = 7, and so forth).

TABLE 7

Distribution of MW and Non-MW Observations across Credit Rating Classes

	Non-MW Firms Pre-404	Non-MW Firms Post-404	MW Firms Pre-404	MW Firms Post-404
S&P Credit Rating				
AAA rating	0.4%	0.4%	0.0%	0.0%
AA rating range	2.9%	2.9%	0.0%	0.0%
A rating range	20.4%	20.4%	6.7%	6.7%
BBB rating range	34.1%	33.4%	13.3%	13.3%
BB rating range	26.2%	27.3%	36.7%	33.3%
B rating range	15.3%	14.9%	40.0%	43.4%
CCC+ or below	0.7%	0.7%	3.3%	3.3%
Summary				
A− or above rating	23.7%	23.7%	6.7%	6.7%
Rating from BBB+ to B−	75.6%	75.6%	90.0%	90.0%
Rating below B−	0.7%	0.7%	3.3%	3.3%

The largest percentage in each column is rounded up or down by 0.1 percent to achieve 100 percent overall.
Table 7 reports the distribution of MW and non-MW observations across various credit rating classes in month $m-1$ ("Pre-404" columns) and in month $m+1$ ("Post-404" columns). Credit rating in month $m-1$ is credit rating at the beginning of the month preceding the month of an initial Section 404 report. Credit rating in month $m+1$ is credit rating at the end of the month following the month of an initial Section 404 report. There are 450 non-MW firms and 30 MW firms in the sample used to construct Table 7. These are the firms that are included in the sample of 577 firms described in Table 1 that also have the necessary ratings data available in the Compustat Ratings File, which reports monthly long-term issuer credit ratings for firms in the S&P 1500.

is that its reporting requirements provide new information to both the private debt market and the public debt market.

VII. SUMMARY

We find marginal support for the hypothesis that, on average, the disclosure of a MW in an initial Section 404 report is associated with an increase in the cost of a firm's publicly traded debt. Further analysis indicates that this average finding does not hold equally between firms that are monitored and firms that are not monitored by credit rating agencies and/or banks. Specifically, we find that the increased cost of debt associated with MW disclosure is more pronounced for firms that are not monitored than for firms that are monitored. Consistent with the argument that banks are effective delegated monitors for the debt market, additional analysis indicates that the effect of bank monitoring appears to be the primary driver of these monitoring results.

When we compare our results to the corresponding cost of equity results in Ashbaugh-Skaife et al. (2009), our additional tests of the impact of monitoring allow us to offer several suggestions for future research. Specifically, we suggest comparing and contrasting the effect of various monitoring agents on (1) the change in a firm's cost of equity and the disclosure of a MW, and (2) the change in a firm's cost of debt and the disclosure of a MW.

Our study tests for a potential benefit of Sarbanes-Oxley, namely, new information provided to the public debt market in a Section 404 report. Testing for the existence of the benefits of Sarbanes-Oxley is important because many corporate executives claim that the costs of Sarbanes-Oxley outweigh the benefits. For example, in an early 2006 Financial Executives Institute survey, more than 80 percent of respondents state that the costs of Sarbanes-Oxley outweigh the benefits (O'Sullivan 2006). Regarding this debate, our results indicate that, on average, a firm's credit spread on its publicly traded debt increases if it discloses a MW in an initial Section 404 report. This finding suggests that one benefit of Section 404 reporting is that it can provide new information to the public debt market.[35] We hasten to add that this conclusion does not consider whether the costs of complying with Section 404 reporting requirements outweigh the benefits.

We conclude with a discussion of sample size and model fit. Regarding sample size, the number of MW firms in some sample cells is quite small. Thus, we acknowledge that some of our inferences are based on small sample cell sizes. However, we believe that these inferences are reliable given that our population of interest in these cells is relatively small to begin with. For example, in our joint tests of monitoring effects, there are five firms in the cell "MW firms monitored by banks but not by credit rating agencies." This small sample size is mainly a function of the relatively small population of firms that satisfy all of the following six conditions: (1) are an accelerated filer, (2) report a MW in an initial Section 404 report, (3) do not report a prior Section 302 MW, (4) have publicly traded debt, (5) are not rated by S&P or Moody's, and (6) have a new or revised bank loan agreement between 1988 and 2005. Regarding model fit, we use changes tests that include variables that attempt to control for factors and events, other than the information in an initial Section 404 report, that could influence the post-report change in a firm's credit spread. However, given the difficulty of identifying and accurately measuring all such factors and events, we acknowledge that our models could be missing some relevant control variables.

REFERENCES

Altman, E. I. 1968. Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance* 23 (4): 589–609.

Altman, E. I., A. Gande, and A. Saunders. 2010. Bank debt versus bond debt: Evidence from secondary market prices. *Journal of Money, Credit and Banking* 42 (4): 755–767.

Ashbaugh-Skaife, H., D. W. Collins, and W. R. Kinney, 2007. The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics* 44 (1–2): 166–192.

Ashbaugh-Skaife, H., D. W. Collins, W. R. Kinney, Jr., and R. LaFond. 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity. *Journal of Accounting Research* 47 (1): 1–43.

Begley, J., J. Ming, and S. Watts. 1996. Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies* 1 (4): 267–284.

Beneish, M. D., M. B. Billings, and L. D. Hodder. 2008. Internal control weaknesses and information uncertainty. *The Accounting Review* 83 (3): 665–704.

Bharath, S. T., J. Sunder, and S. V. Sunder. 2008. Accounting quality and debt contracting. *The Accounting Review* 83 (1): 1–28.

Bhojraj, S., and P. Sengupta. 2003. Effects of corporate governance on bond ratings and yields: The role of institutional investors and outside directors. *Journal of Business* 76 (3): 455–476.

Botosan, C. A., and M. Plumlee. 2005. Assessing alternative proxies for the expected risk premium. *The Accounting Review* 80 (1): 21–53.

Choi, J., and M. P. Richardson. 2010. The volatility of the firm's assets. Working paper, University of Illinois at Urbana–Champaign and New York University. Available at: SSRN: http://ssrn.com/abstract=1359368.

---

[35] As noted previously, consistent with this inference, concurrent research by Kim et al. (2011) and Costello and Wittenberg-Moerman (2011) finds evidence suggesting that one benefit of Sarbanes-Oxley is that its reporting requirements provide new information to the private (i.e., bank loan) debt market.

Costello, A., and R. Wittenberg-Moerman. 2011. The impact of financial reporting quality on debt contracting: Evidence from internal control weakness reports. *Journal of Accounting Research* 49 (1): 97–136.

Crabtree, A., J. J. Maher, and D. Wang. 2009. Corporate internal control disclosures and credit ratings. Working paper, University of Nebraska-Lincoln, Virginia Tech, and Texas A&M University.

Datta, S., M. Iskandar-Datta, and A. Patel. 2000. Do bank relationships matter in public debt offerings? *Journal of Applied Corporate Finance* 12 (4): 120–127.

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violations and manipulation of accruals. *Journal of Accounting and Economics* 17 (1–2): 145–176.

Diamond, D. 1984. Financial intermediation and delegated monitoring. *The Review of Economic Studies* 51 (3): 393–414.

Doyle, J., W. Ge, and S. McVay. 2007. Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics* 44 (1–2): 193–223.

Duffee, G. R. 1998. The relation between treasury yields and corporate bond yield spreads. *Journal of Finance* 53 (6): 2225–2241.

Elbannan, M. A. 2009. Quality of internal control over financial reporting, corporate governance and credit ratings. *International Journal of Disclosure and Governance* 6 (2): 127–149.

Fama, E. 1985. What's different about banks? *Journal of Monetary Economics* 15 (1): 29–39.

Federal Reserve Board. 2006. Federal Reserve Statistical Release H.15 (Treasury Yields). Available at: http://www.federalreserve.gov/RELEASES/H15/data.htm.

Ge, W., and S. McVay. 2005. The disclosure of material weaknesses in internal control after the Sarbanes-Oxley Act. *Accounting Horizons* 19 (3): 137–158.

Hogan, C. E., and M. S. Wilkins. 2008. Evidence on the audit risk model: Do auditors increase audit fees in the presence of internal control deficiencies? *Contemporary Accounting Research* 25 (1): 219–242.

James, C. M. 1987. Some evidence on the uniqueness of bank loans. *Journal of Financial Economics* 19 (2): 217–235.

Kim, J., B. Y. Song, and L. Zhang. 2011. Internal control weakness and bank loan contracting: Evidence from SOX Section 404 disclosures. *The Accounting Review* 86 (4): 1157–1188.

Lambert, R. A., C. Leuz, and R. E. Verrecchia. 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45 (2): 385–420.

Longstaff, F. A., and E. S. Schwartz. 1995. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance* 50 (3): 789–819.

Mansi, S. A., W. F. Maxwell, and D. P. Miller. 2004. Does auditor quality and tenure matter to investors? Evidence from the bond market. *Journal of Accounting Research* 42 (4): 755–793.

Mergent Inc. 2005. *Mergent Bond Record.* New York, NY: Moody's Investor Services, Inc.

Moody's Investors Service, Inc. 2004. *Moody's Special Comment: Section 404 Reports on Internal Control: Impact on Ratings Will Depend on Nature of Material Weaknesses Reported.* New York, NY: Moody's Investor Services, Inc.

Moody's Investors Service, Inc. 2005a. *Moody's Special Comment: Credit Implications for Companies That Fail to File Internal Control Reports on Time.* New York, NY: Moody's Investor Services, Inc.

Moody's Investors Service, Inc. 2005b. *Moody's Special Comment: Section 404 Reporting on Internal Control: Our Early Experience.* New York, NY: Moody's Investor Services, Inc.

National Association of Securities Dealers (NASD). 2004. *An Overview of the Regulation of the Bond Markets.* Testimony of Doug Shulman, President NASD Markets, Services and Information, before the U.S. Senate Committee on Banking, Housing, and Urban Affairs (June 17). Available at: http://banking.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=182c34a9-bedb-4f65-ba0c-75d98f728525.

Neal, R., D. S. Rolph, and C. Morris. 2001. Interest rates and credit spread dynamics. Working paper, Indiana University and City University of Hong Kong. Available at SSRN: http://ssrn.com/abstract=249983.

Ogneva, M., K. Raghunandan, and K. R. Subramanyam. 2007. Internal control weakness and implied cost of equity: Evidence from SOX Section 404 disclosures. *The Accounting Review* 82 (5): 1255–1297.

Ohlson, J. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18 (1): 109–131.

O'Sullivan, K. 2006. The case for clarity. *CFO* (September 1): 64–69. Available at: http://www.cfo.com/article.cfm/7851741/c_7873404.

Petersen, M. A., and R. G. Rajan. 1994. The benefits of lending relationships: Evidence from small business data. *Journal of Finance* 49 (1): 3–37.

Public Company Accounting Oversight Board (PCAOB). 2004. *An Audit of Internal Control over Financial Reporting Performed in Conjunction with an Audit of Financial Statements.* Auditing Standard No. 2. Washington, D.C.: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2007. *An Audit of Internal Control over Financial Reporting that is Integrated with an Audit of Financial Statements and Related Independence Rule and Conforming Amendments.* Auditing Standard No. 5. Washington, D.C.: PCAOB.

Securities and Exchange Commission (SEC). 2003. *Final Rule: Management's Report on Internal Control over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports.* Release Nos. 33–8238 and 34–47986; File Nos. S7-40-02 and S7-06-03. Washington, D.C.: SEC.

Securities and Exchange Commission (SEC). 2004. *Order under Section 36 of the Securities Exchange Act of 1934 Granting an Exemption from Specified Provisions of Exchange Act Rules 13a-1 and 15d-1.* Release No. 34–50754. Washington, D.C.: SEC.

Securities and Exchange Commission (SEC). 2005. *Self-Regulatory Organizations: National Association of Securities Dealers, Inc.: Notice of Filing of Proposed Rule Change Relating to Dissemination of Information on TRACE-Eligible Securities Transactions.* Release No. 34–52700; File No. SR-NASD-2005-120. Washington, D.C.: SEC.

Securities Industry and Financial Markets Association (SIFMA). 2009. Investing in bonds. Available at: http://www.investinginbonds.com/learnmore.asp?catid=5&subcatid=18&id=174.

Sengupta, P. 1998. Corporate disclosure quality and the cost of debt. *The Accounting Review* 73 (4): 459–474.

Standard & Poor's (S&P). 2004. *Corporate Rating Criteria.* New York, NY: Standard & Poor's.

APPENDIX A

Data Sources for Test Variables

Yield = the mean YTM for all TRACE-reported trades in a firm's debt issues over a specified time period. Data Source: WRDS TRACE data;

Spread = the mean credit spread for all TRACE-reported trades in a firm's debt issues over a specified time period. We measure the credit spread for each trade as YTM on the relevant debt issue minus YTM on a matched-by-maturity Treasury security. Data Sources: WRDS TRACE data and Treasury yield data;

MW = 1 if a firm reports a material weakness in an initial Section 404 report, and 0 if it does not report a material weakness in an initial Section 404 report. Data Source: Audit Analytics;

Rated = 1 if a firm's debt is rated by S&P and/or Moody's, and 0 otherwise. Data Sources: Compustat (S&P rating) and Mergent Bond Record (Moody's rating);

Bank = 1 if a firm has a new or revised bank loan agreement in any year from 1988 to 2005, and 0 otherwise. Data Source: DealScan;

DefaultRisk = the score from the bankruptcy model of Ohlson (1980). Higher values of *DefaultRisk* indicate a higher probability of default risk. Data Source: Compustat;

Size = the log of total assets. Data Source: Compustat;

CFO = cash flow from operations scaled by total assets. Data Source: Compustat;

Leverage = long-term debt scaled by total assets. Data Source: Compustat;

CFVol = the standard deviation over the last 20 quarters of cash flow from operations scaled by total assets. Data Source: Compustat;

TrsYTM = the mean of all "YTM on a matched-by-maturity Treasury security" used to calculate a firm's *Spread* over a specified time period. Data Source: Treasury yield data; and

Restructure (*WriteDown*; *M&A; Restate*) = an indicator variable equal to 1 if restructuring (writedown; M&A; restatement) activity occurs in a given period, and 0 otherwise. Data Sources: Compustat and Audit Analytics.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1157–1188

American Accounting Association
DOI: 10.2308/accr-10036

Internal Control Weakness and Bank Loan Contracting: Evidence from SOX Section 404 Disclosures

Jeong-Bon Kim
City University of Hong Kong

Byron Y. Song
Concordia University

Liandong Zhang
City University of Hong Kong

ABSTRACT: Using a sample of borrowing firms that disclosed internal control weaknesses (ICW) under Section 404 of the Sarbanes-Oxley Act, this study compares various features of loan contracts between firms with ICW and those without ICW. Our results show the following. First, the loan spread is higher for ICW firms than for non-ICW firms by about 28 basis points, after controlling for other known determinants of loan contract terms. Second, firms with more severe, company-level ICW pay significantly higher loan rates than those with less severe, account-level ICW. Third, lenders impose tighter nonprice terms on firms with ICW than on those without ICW. Fourth, fewer lenders are attracted to loan contracts involving firms with ICW. Finally, our within-firm analyses show that banks increase loan rates charged to ICW firms after their disclosure of internal control problems and that banks reduce loan rates after firms remediate previously reported ICW.

Keywords: *internal control weaknesses; loan contracting; Sarbanes-Oxley Act (SOX).*

JEL Classifications: *G21; G32; K22; M41.*

We are especially grateful to Steven Kachelmeier (the senior editor) and two anonymous referees for their many insightful and constructive suggestions. We appreciate helpful comments from Leslie Hodder (AAA discussant), Bin Ke, Mozaffar Khan (CAPANA discussant), Yinghua Li, Steve Penman, Annie Qiu, Haina Shi, Shyam V. Sunder, Zheng Wang, Cheong Yi, Sanjian Zhang (CAAA discussant), and participants of research workshops at Concordia University, City University of Hong Kong, Shanghai Jiaotong University, The Hong Kong Polytechnic University, Sun Yat-Sen University, and the 2008 Annual Conference of CAAA, the 2009 Annual Research Conference of CAPANA, the 2009 AAA Annual Meeting. Any errors are our own. The major part of this research was completed while Professor Kim worked as Canada Research Chair in Corporate Governance and Financial Reporting at Concordia University. Professor Kim acknowledges partial financial support for this project from the Social Sciences and Humanities Research Council of Canada via the Canada Research Chair program. Professors Song and Zhang acknowledge partial financial support for this project from the Faculty Research Development Program (FRDP), John Molson School of Business, Concordia University.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: November 2009
Accepted: December 2010
Published Online: April 2011

Data Availability: *All data are available from sources identified in the study.*

I. INTRODUCTION

In response to a series of corporate failures and accounting scandals that started with the Enron debacle and the subsequent Andersen collapse, the U.S. Congress passed the Sarbanes-Oxley Act (SOX) in July 2002. Section 404 of SOX (SOX 404) requires company management to assess the effectiveness of internal controls over financial reporting and to provide periodic auditor-attested evaluations of internal control effectiveness. The major rationale behind this requirement is that lack of adequate internal controls leads to poor financial reporting and increases information risk, which in turn increases the cost of external financing (e.g., U.S. House of Representatives 2005).

Several studies have examined the impact of material internal control weakness (ICW) on information uncertainty and the cost of capital from the perspective of equity stakeholders, and the evidence provided is mixed (e.g., Ogneva et al. 2007; Beneish et al. 2008; Hammersley et al. 2008; Ashbaugh-Skaife et al. 2009). For example, using a sample of SOX 404 filers with the U.S. Securities and Exchange Commission (SEC), Ogneva et al. (2007) find no significant differences in their implied cost of equity estimates between ICW firms and non-ICW firms. In contrast, Ashbaugh-Skaife et al. (2009) report that ICW firms exhibit significantly higher beta risk and cost of equity capital than non-ICW firms, suggesting that information risk arising from ICW is priced in the stock market. However, previous research has paid little attention to the economic consequences of internal control quality from the perspective of credit stakeholders. To help fill this void, our study aims to provide systematic evidence on the effects of ICW on bank loan contracting. Specifically, our analyses focus on the manner in which ICW impacts various features of loan contracts, such as loan spreads, collateral requirements, restrictive covenants, and the number of lenders for a loan.

We are motivated to examine the loan contracting consequences of ICW for several reasons. First, bank loans are a major source of external financing in the United States and most other countries (e.g., Graham et al. 2008).[1] Thus, it is economically important to investigate whether internal control quality is associated with the cost of bank loans. This investigation has the potential to provide evidence on the benefit of maintaining and reporting internal control effectiveness according to SOX 404. Second, measurement errors inherent in the implied cost of equity estimates can lead to the inconclusive results of prior research on the relation between ICW and the cost of capital.[2] In contrast, the cost of private debt is observed directly through loan contract terms and is subject to fewer measurement errors. This allows us to make a cleaner inference about the impact of ICW on the cost of external financing. Finally, multifaceted features of loan contracts provide a unique opportunity to assess not only the direct cost of ICW (i.e., higher loan spreads), but also the associated indirect cost (i.e., tighter nonprice terms).[3]

---

[1] Since 1980, bank loans have consistently accounted for more than 50 percent of total debt financing in the United States (see Graham et al. 2008). The volume of syndicated loans reached $1.69 trillion in 2007, according to the DealScan database.

[2] Various implied cost of equity measures are generally inferred from the current stock price, based on some valuation models. Thus, the accuracy of these measures depends critically on the validity of the assumptions and structures of the valuation models used. Moreover, some key inputs to the valuation models can also suffer from measurement errors. For example, it is widely recognized in the literature that analyst earnings forecasts, a proxy for expected future earnings in these models, are biased upward.

[3] The restrictive covenants can also limit flexibility in investment decisions or cause the borrower to abandon profitable investment projects to comply with them, which increases the indirect cost of debt.

To provide evidence of the impact of ICW on various features of loan contracts, we construct a sample of 3,164 loan facility-years for borrowers that filed SOX 404 disclosures with the SEC during 2005–2009. We then compare various features of loan contracts with ICW borrowers with those with non-ICW borrowers, after controlling for borrower- and loan-specific characteristics deemed to affect the contract terms. Briefly, our main results show the following.

First, we find that the loan spread is higher for ICW firms than for non-ICW firms by about 28 basis points (bps) in the full-model regressions, after controlling for other factors influencing loan terms, default risk, credit quality, and internal control quality. This finding is consistent with the notion that banks take into account internal control over financial reporting when setting the price term of loan contracts. That is, banks view ICW as an incrementally significant information risk factor, above and beyond traditional credit risk factors, that increases pre-contract information uncertainty, as well as post-contract monitoring and re-contracting costs. Our results suggest that the quality of the financial reporting system plays an important role in private debt contracting and that information risk incurred by weak internal controls cannot be removed by lenders' access to borrowers' inside information.

Second, we find that the nature or severity of material weakness in internal control matters in loan contracting. Specifically, we find that borrowers with more severe, company-level ICW (e.g., organizational control or governance flaws) pay higher loan rates than those with less severe, account-level ICW (e.g., inventory recording flaws and lease accounting problems). This finding suggests that lenders are able to differentiate more severe ICW problems from less severe ones when designing loan contracts. To some extent, it implies that the comparative advantages that banks have in accessing inside information help them reduce, but not completely eliminate, the information risk associated with account-level control problems, but lenders have more difficulty obtaining useful inside information that can help mitigate information risk associated with company-level control problems.

Third, we find that lenders impose tighter nonprice terms on ICW borrowers than on non-ICW borrowers. In particular, we find that the likelihood of a loan being secured by collateral is higher for ICW borrowers than for non-ICW borrowers. We also provide evidence that collateral and restrictive covenants are used more intensively in loan contracts involving borrowers with company-level control problems than in those involving borrowers with account-level control problems. Fourth, we find that the number of lenders in each loan is smaller for loans to ICW borrowers than for those to non-ICW borrowers, consistent with the theory that information asymmetries between the borrower and potential lenders attract fewer lenders in a loan syndicate (e.g., Sufi 2007).

Fourth, using a sample of firms that initiate similar loans both before and after their first-time internal control disclosures under SOX 404, we conduct within-firm analyses and find that the loan spreads charged to ICW firms significantly increase after SOX 404 disclosures, while the loan spreads charged to non-ICW firms significantly decrease after SOX 404 disclosures. This suggests that SOX 404 disclosures provide new information to the private debt market regarding a borrower's information risk. Moreover, we show that firms enjoy a meaningful reduction in the cost of bank loans after they remediate previously reported ICW problems, consistent with the equity market evidence documented by Ashbaugh-Skaife et al. (2009).

Finally, we find that the number of financial covenants imposed on loans to *both* ICW and non-ICW firms decreases after SOX 404 disclosures. In a concurrent study, Costello and Wittenberg-Moerman (2011) find that the number of financial covenants decreases after a firm discloses ICW under SOX Section 302 (SOX 302). The authors interpret this result as evidence that ICW disclosure *per se* leads to decreased use of financial covenants. However, because Costello and Wittenberg-Moerman (2011) include only ICW firms in their sample, their within-firm analysis seems to be missing a control condition, which could limit the inferences they may draw from their analysis.[4] We conjecture that the decreased use of financial covenants for both ICW and non-ICW firms could be driven, at least partially, by a general regulatory effect of SOX.[5]

This study contributes to the existing literature in at least three ways. First, our study adds to the literature on the economic consequences of the quality of the financial reporting system in the debt market. Unlike studies of debt contracting that rely on researchers' self-constructed reporting quality proxies (e.g., Bharath et al. 2008), we use auditor-attested evaluations of material weakness in internal control, a direct signal of a low-quality financial reporting system. Our results suggest that the quality of the financial reporting system plays an important role in the private debt market, even though the lenders in this market have superior information processing capabilities and better access to borrowers' inside information. Our study is related to Bharath et al. (2008) and Graham et al. (2008), who show that accrual quality and financial restatement are associated with the cost of private debt. We show that our results remain the same after controlling for accrual quality and financial restatement. This finding suggests that our results are largely driven by the dimensions of a firm's information system quality that is incremental to the accounting quality measures used in prior research.[6]

The concurrent studies by Dhaliwal et al. (2011) and Costello and Wittenberg-Moerman (2011) also examine the relation between ICW and cost of debt. Dhaliwal et al. (2011) find that the cost of debt increases after the disclosure of a material weakness in an initial SOX 404 report, which is consistent with our within-firm analysis. A key difference between our study and Dhaliwal et al. (2011) is that their analysis focuses on the credit spread in the secondary bond market (i.e., public debt market), whereas our approach focuses on the cost of debt in the primary bank loan market (i.e., private debt market).[7] In addition, Dhaliwal et al. (2011) use change analysis, while our study conducts both cross-sectional and change analyses. Costello and Wittenberg-Moerman (2011) examine the impact of ICW disclosures on the cost of bank loans and report somewhat similar findings to our within-firm analysis. An important feature that distinguishes their study from ours is that their focus is on unaudited SOX 302 disclosures, whereas ours is on auditor-attested SOX 404 disclosures. We argue that, compared to SOX 302 disclosures, auditor-attested SOX 404 disclosures are more reliable indicators of a firm's financial reporting system quality. Moreover, as their study investigates only the temporal changes in loan terms for ICW firms, they are unable to provide evidence on the differences in various loan features between loans to ICW firms and loans to non-ICW firms.

Second, our study contributes to the loan contracting literature, as well. Evidence reported in this study suggests that information risk associated with ICW is a priced risk factor that is distinct from traditional credit risk factors. Our results also show that banks not only incorporate ICW information into loan contract terms, but they are also able to differentiate between company- and account-level ICW problems when designing loan contracts and monitoring post-contract credit quality.

---

[4] See Section VI for detailed discussions.

[5] For example, the improved general information environment and governance by SOX could reduce the need for bank monitoring through financial covenants. In our tests, we do find that the number of financial covenants decreases more (significant at the 10 percent level) for ICW firms than for non-ICW firms. However, this could be interpreted as evidence that the general regulatory impact of SOX on ICW firms is more pronounced because of their higher *ex ante* information asymmetry.

[6] Section II provides in-depth discussions on the incremental usefulness of our study in establishing the empirical link between accounting quality (or, more generally, information risk) and cost of debt.

[7] As discussed further in Section II, unlike investors in the bond market, banks have comparative advantages in accessing borrowers' inside information. For example, Fama (1985) argues that banks have superior information due to their access to information from an organization's decision process not otherwise publicly available. In contrast, public debt holders (e.g., investors of publicly traded bonds) rely mostly on publicly available information.

Finally, our study provides new evidence on the benefits of SOX 404 internal control disclosure in the private debt market, and thus contributes to the debate on the costs and benefits of SOX 404 compliance. The combined empirical evidence in Ashbaugh-Skaife et al. (2009), Dhaliwal et al. (2011), and our study suggests that SOX 404 disclosure is informative to stakeholders in the equity market, the public debt market, and the private debt market.

Section II develops research hypotheses. Section III presents the empirical model for hypothesis testing. Section IV describes the sample and data sources, presents descriptive statistics, and discusses univariate results. Section V reports the results of multivariate tests on the impact of ICW on price and nonprice terms and the number of lenders. Section VI reports the results of within-firm analyses and various robustness checks. Section VII concludes.

II. MOTIVATION AND HYPOTHESES

Duffie and Lando (2001) develop a theory that information risk faced by lenders is incremental to borrower default risk. Their analysis indicates that credit suppliers require compensation not only for fundamental credit risk, but also for imperfect information-related risk. Easley and O'Hara (2004) and Lambert et al. (2007) develop similar theoretical predictions on the relation between information risk and cost of capital. Several recent empirical studies provide evidence supporting the above theories. For example, using accrual quality as an inverse measure of information risk, Francis et al. (2005) show that accrual quality is negatively related to a firm's average cost of debt. Similarly, Bharath et al. (2008) find that banks charge lower loan spreads and impose less stringent nonprice terms for firms with higher accrual quality. Finally, Graham et al. (2008) show that borrowers that restate their financial statements tend to be given unfavorable loan contract terms.

Our study extends the above research by examining the relation between ICW and bank loan contracting. We posit that weak internal controls introduce both intentional reporting biases and unintentional accounting errors into the financial reporting process, thereby increasing information risk and the cost of bank loans (Public Company Accounting Oversight Board [PCAOB] 2007). Despite the findings of previous research that ICW is related to accrual quality (e.g., Doyle et al. 2007b; Ashbaugh-Skaife et al. 2008) and that accrual quality is associated with the cost of debt (Bharath et al. 2008), our research is useful for several reasons.

First, the implications from accrual quality research are limited by the potential inaccuracy inherent in the accrual quality measures used; their accuracy relies typically on the validity of (discretionary) accrual models, which have been attracting increasing criticism.[8] Given the problems inherent in accrual quality measures, the empirical link between accounting information quality and cost of debt is far from settled. On the other hand, a SOX 404 auditor-attested evaluation of an internal control system is a relatively less ambiguous accounting quality measure (e.g., Ashbaugh-Skaife et al. 2009). Thus, examining the relation between ICW and the cost of bank loans is a useful exercise to further our understanding of the impact of accounting system quality on the cost of debt.

Second, even if accrual quality is measured without error, such a measure captures only the quality of the bottom line item of a firm's external financial reports. However, there is much more information in a firm's financial reports than just earnings numbers. For example, given their superior information processing capability, banks can overcome the problem of low-quality accruals by carefully examining the information in the notes to the financial statements. In contrast, ICW disclosures reflect the quality of a firm's entire information production system, which will, in turn, affect the quality of a variety of information items in the financial reports, including notes to financial statements and many other non-bottom-line items.

---

[8] See Dechow et al. (2010) for a comprehensive discussion on the problems of the various accrual quality measures.

Third, unlike arm's-length investors in the equity and bond markets, commercial banks and other private lenders have comparative advantages in accessing borrowers' inside information that is useful for mitigating pre-contract information uncertainty and post-contract monitoring problems (Fama 1985; Cole 1998; Rajan 1992; Denis and Mihov 2003; Bharath et al. 2008). As a result, banks may be able to overcome information problems associated with poor-quality financial reporting at a relatively low cost. This implies that the quality of a firm's external financial reporting may be of lesser importance to banks than to widely dispersed equity and/or bondholders. However, the proxies used in the literature capture only the quality of externally reported earnings (e.g., accrual quality, restatements). In contrast, a firm's internal control system can affect the quality of its inside information, as well as information in externally oriented financial reports. For example, Feng et al. (2009) find that management earnings guidance is less accurate for firms with ICW, suggesting that ICW affects the quality of internal reports used by managers. The authors argue that even insiders cannot overcome problems arising from ineffective internal controls. Thus, we argue that ICW is a comprehensive information quality indicator that is able to capture the quality of both inside and publicly reported accounting information.

Overall, we believe that ICW is a more comprehensive and accurate indicator of a firm's accounting information risk and that establishing the link between ICW and the cost of bank loans can provide more convincing evidence for the pricing of accounting information risk in the private debt market. Moreover, documenting the benefits of effective internal controls and internal control disclosures is interesting in its own right, given the substantial cost of SOX 404 compliance. To provide empirical evidence on the role of internal control quality in loan pricing, we test the following hypothesis in alternative form:

H1: Controlling for other determinants, loan spreads are higher for ICW borrowers than for non-ICW borrowers.

Not all internal control material weaknesses have the same nature and potential severity. Moody's Investors Service (2004, 2006, 2007; hereafter, Moody's) classifies the various material weaknesses into two types, account-level material weaknesses (Category A) and company-level material weaknesses (Category B). Account-level material weaknesses relate to controls over specific account balances or transaction-level processes, either of which is relatively "auditable." Moody's is less concerned about this type of material weakness and does not expect to take any rating action for firms reporting it. Company-level material weaknesses relate to the overall control environment or financial reporting process, and they have a pervasive effect on a company's financial reporting. Moody's contends that company-level material weakness calls into question not only management's ability to prepare accurate financial reports, but also its ability to control the business. Thus, Moody's is more concerned about company-level material weakness and more likely to take rating action for firms with company-level control problems.

Consistent with Moody's arguments, previous research documents that account- and company-level material weaknesses have different impacts. For example, Doyle et al. (2007b) find that the presence of company-level weakness results in lower accruals quality, but find no significant relation between the presence of account-level weakness and accrual quality. This evidence suggests that company-level weakness hampers the quality of financial reporting to a greater extent than does account-level weakness. To examine whether the nature or severity of ICW matters in loan pricing, we test the second hypothesis in alternative form:

H2: Controlling for other determinants, loan spreads are higher for borrowers with company-level ICW than for those with account-level ICW.

Bank loan contracts include not only the price term (i.e., loan spread), but also various nonprice terms, such as collateral requirements and restrictive covenants. Lenders use these nonprice terms (as well as the price term) when designing loan contracts in an attempt to mitigate (1) agency costs of debt arising from agency conflicts between shareholders and debtholders (Smith and Warner 1979), (2) information problems faced by lenders (Graham et al. 2008; Kim et al. 2010), and (3) potential conflicts between lenders and borrowers (Vasvari 2008). Extant evidence shows that lenders are more likely to require collateral for borrowers with opaque information (e.g., Berger and Udell 1990; Rajan and Winston 1995; Jimenez et al. 2006). We therefore expect that banks are more likely to require borrowers with ICW problems to pledge collateral on their loans. This leads to our third hypothesis:

> **H3:** Controlling for other determinants, the likelihood of loans being secured by collateral is higher for ICW borrowers than for non-ICW borrowers.

The debt covenant literature indicates that lenders use loan covenants to improve the *ex post* monitoring of changes in credit quality, although the use of covenants also incurs costs of reduced flexibility for the borrower (Smith and Warner 1979; Rajan and Winston 1995; Bradley and Roberts 2004). In particular, Rajan and Winston (1995) show that the inclusion of restrictive covenants in loan contracts provides lenders with a strong incentive to monitor more closely credit quality subsequent to loan initiations. To the extent that lenders are likely to have higher incentives to monitor the post-contract performance of borrowers with poor reporting quality, we expect that restrictive covenants are more intensively used in loans involving borrowers with ICW problems than in those involving borrowers without such problems. This leads to our fourth hypothesis:

> **H4:** Controlling for other determinants, the use of restrictive covenants in loan contracts is more intensive and prevalent for ICW borrowers than for non-ICW borrowers.

The syndicate loan literature shows that fewer lenders are involved in a loan deal when borrowers have information problems. This occurs because a syndicate structure with fewer lenders helps minimize free-rider problems in private information gathering and facilitates *ex post* monitoring (Dennis and Mullineaux 2000; Sufi 2007). Thus, we expect the number of lenders to be smaller for loan contracts with ICW borrowers than for loan contracts with non-ICW borrowers. This leads to our final hypothesis:

> **H5:** Controlling for other determinants, the number of lenders in a loan deal is smaller for ICW borrowers than for non-ICW borrowers.

III. EMPIRICAL MODEL

To evaluate the impact of ICW on various features of loan contracting, we specify the following regression model:

$$\begin{aligned} Loan\_Feature = {} & \alpha_0 + a_1 Weak + \alpha_2 Loan\text{-}specific\_Control \\ & + a_3 Borrower\text{-}specific\_Control + a_4 Economy\text{-}wide\_Control \\ & + (Year\_Indicators) + (Industry\_Indicators) + Error, \end{aligned} \tag{1}$$

where, for each loan facility and borrowing firm, all variables are as defined in Appendix A. The dependent variable, *Loan_Feature*, refers to one of the following features of a loan contract: (1) the price term, namely, the drawn all-in spread (*AIS*); (2) the nonprice terms, including the likelihood of the loan being secured by collateral (*DSecured*), the number of financial covenants included in each loan (*FinCovIndex*), and the number of general covenants included in each loan (*GenCovIndex*); and (3) the number of lenders involved in the loan deal (*NLenders*).

The variable *AIS* proxies for the interest cost of borrowing and is measured by the drawn all-in spread (plus the upfront fee and annual fee, if any) in basis points in excess of the benchmark rate, that is, the London Inter-Bank Offer Rate (LIBOR). The test variable *Weak* is an indicator variable that equals 1 if the auditor concludes in the SOX 404 report filed with the SEC that the borrower's internal control over financial reporting is ineffective for the fiscal year, and 0 otherwise. When *AIS* is used as the dependent variable, that is, *Loan_Feature* in Equation (1), the coefficient on *Weak* captures the difference in the loan spread charged to borrowers with ICW problems *vis-à-vis* those without such problems.[9] Therefore, H1 implies $\alpha_1 > 0$. Similarly, H3 (H4) is supported if we observe $\alpha_1 > 0$ when the dependent variable is *DSecured* (*FinCovIndex* or *GenCovIndex*). H5 implies $\alpha_1 < 0$ when the dependent variable is *NLenders*.

To test whether company-level ICW differs systematically from account-level ICW in terms of its impact on loan contracting features (H2),[10] we also estimate Equation (1) after replacing *Weak* by two indicator variables, *Company_ICW* and *Account_ICW*, which equal 1 for borrowers that report any company-level ICW (and borrowers with incomplete internal control assessment)[11] and for those that report any account-level weakness, respectively, and 0 otherwise. H2 is supported if the coefficient on *Company_ICW* is greater than that on *Account_ICW*.

The loan contracting literature shows that several loan-specific characteristics are related to the interest cost of borrowing (e.g., Strahan 1999; Dennis et al. 2000; Bharath et al. 2008; Graham et al. 2008; Vasvari 2008; Kim et al. 2010). We include in Equation (1) a set of loan-level control variables, that is, *Log_Maturity*, *Log_Loan_Size*, *Log_NLenders*, and *Performance_Pricing*, when examining the effect of ICW on loan spreads. The variable *Log_Maturity* is the natural log of loan maturity in months, and *Log_Loan_Size* is measured by the natural log of the dollar amount of each loan facility given to a borrower. Previous research finds that lenders charge lower interest rates for shorter-maturity loans and larger loan facilities (e.g., Graham et al. 2008). Therefore, we expect a positive coefficient for *Log_Maturity* and a negative coefficient for *Log_Loan_Size*. The variable *Log_NLenders* is the natural log of the number of lenders in a loan deal, and *Performance_Pricing* is an indicator variable that equals 1 for loans with performance pricing provisions, and 0 otherwise. We expect loan contracts involving larger numbers of lenders and performance pricing provisions to have lower interest rates.

We also control for a set of borrower-specific variables that are known to affect credit quality and, thus, loan contract terms: *Size*, *Leverage*, *MB*, *Profitability*, *Tangibility*, and *CashVol*. The variables *Size* and *Leverage* are measured by the natural log of total assets and the ratio of long-term debt to total assets, respectively. We expect *Size* (*Leverage*) to be positively (negatively) related to credit quality. The variable *MB* is the market value of equity plus the book value of debt divided by the book value of total assets. To the extent that *MB* proxies for a borrower's growth potential, *MB* is likely to be positively associated with credit quality. However, growing firms are often faced with high risk. In such a case, *MB* is likely to be inversely associated with credit quality. The variable *Profitability* refers to earnings before interest, taxes, depreciation, and amortization (EBITDA) divided by total assets; *Tangibility* is the ratio of property, plant, and equipment (PP&E) to total assets; and *CashVol* refers to the standard deviation of quarterly cash flows from operations (scaled by yearly total assets) over the past five fiscal years. We expect *Profitability* and *Tangibility* to be positively associated with credit quality, and *CashVol* to be inversely associated with credit quality.

---

[9] We define a borrower as a non-ICW firm if its external auditor concludes that the firm maintains, in all material respects, effective internal control over financial reporting in the annual report.

[10] In our study, company-level ICW includes problems related to inadequate disclosure control, an ineffective or understaffed audit committee, lack of senior management competency and tone, ineffective internal audit function, ineffective personnel, and the segregation of duties, while account-level control weaknesses include all other weaknesses.

[11] Moody's Investors Service (2006, 2007) argues that the inability to complete the Section 404 report is itself a company-level control weakness that generally merits rating committee consideration. Thus, we view firms that did not finish internal control assessments as having company-level weaknesses. In our sample, only two firms fall into this category. Excluding these two firms from the sample does not change our empirical results.

To control for borrower default risk, we construct a comprehensive measure of default risk incorporating three default risk proxies that have been widely used in the literature: Altman's (1968) z-score, Ohlson's (1980) O-score, and Shumway's (2001) probability of bankruptcy. Specifically, we apply principal component analysis to the three proxies and extract the first principal component, which is defined as:[12]

$$\begin{aligned} \textit{Default\_Risk} = {} & -0.7999 \times \textit{z-score} + 0.8948 \times \textit{O-score} + 0.6239 \\ & \times \textit{Probability of Bankruptcy}. \end{aligned}$$

We then rank it in each fiscal year and use the decile rank value (*RDefault_Risk*) as our default risk measure in our regression analyses. This ranked measure is constructed in such a way that a higher value of *RDefault_Risk* represents a higher default risk.

Hammersley et al. (2008) show that ICW disclosures are often accompanied by restatements. Further, Graham et al. (2008) document that loans initiated after restatements have significantly higher spreads, shorter maturities, higher likelihoods of loans being secured by collateral, and more covenant restrictions. Thus, restatements may be an important confounding factor in our study.[13] To ensure that our results are not driven by the effect of financial restatements, we include in our model an indicator variable, *Restate*, which equals 1 if the firm restates its financial statements in the year of ICW disclosure. Moreover, Bharath et al. (2008) find that accrual quality affects the terms of bank loan contracts. To alleviate the concern that our results may simply pick up the effects of accrual quality on the cost of debt, we also control for accrual quality (*AbsAccr*) in our regression model. The variable *AbsAccr* is estimated using the modified Jones model (Dechow et al. 1995), allowing for accounting conservatism (Ball and Shivakumar 2006).

Previous research has identified additional firm characteristics that are possible determinants of ICW (Ashbaugh-Skaife et al. 2007; Doyle et al. 2007a). To the extent that these variables are correlated with both ICW and loan contract terms, their omission could create a problem of correlated omitted variables. To avoid this problem, we also control for the following additional firm characteristics: auditor quality (*Big4*), firm age (*Age*), the number of the firm's business segments (*NSegment*), inventory holdings (*Inventory*), a firm's involvement in mergers and acquisitions (*M&A*), organizational change (*Restructure*), the existence of foreign sales (*Foreign*), sales growth (*Sale_Growth*), and the presence of aggregate loss (*Loss*).[14]

In addition, we include two economy-wide variables, *Credit_Spread* and *Term_Spread*, to control for the potential effects of macroeconomic conditions on loan contracting. Here, *Credit_Spread* is the difference in yield between BAA- and AAA-rated corporate bonds, while *Term_Spread* is the difference in yield between ten-year and two-year U.S. Treasury bonds.

---

[12] Here, z-score = 1.2 × (working capital/total assets) + 1.4 × (retained earnings/total assets) + 3.3 × (EBIT/total assets) + 0.6 × (market value of equity/total liabilities) + 0.999 × (sales/total assets); O-score = −1.32 − 0.407 × log(total assets) + 6.03 × (total liabilities/total assets) − 1.43 × (working capital/total assets + 0.076 × (current liabilities/current assets) − 1.72 × (1 if total liabilities > total assets, 0 otherwise) − 2.37 × (net income/total assets) − 1.83 × (operating income before depreciation/total liabilities) + 0.285 × (1 if net income is negative for the last two years, 0 otherwise) − 0.521 × ((net income_t − net income_{t-1})/(|net income_t|+|net income_{t-1}|)); and Probability of Bankruptcy = $e^w/(1+e^w)$, where w = −13.303 − 1.982 × (net income/total assets) + 3.593 × (total liabilities/total assets) − 0.467 × log (firm's market capitalization/total market capitalization) − 1.809 × excess stock return over the value-weighted index return in the past 12 months + 5.791 × stock return volatility in the past 12 months.

[13] Note, however, that a restatement is in itself a reflection of ineffective internal controls.

[14] Refer to Appendix A for detailed definitions of these variables.

Finally, we include *Loan_Type_Indicators* and *Loan_Purpose_Indicators* to control for potential differences in the price and nonprice terms of loan contracts associated with the different types and purposes of loans. We also include *Industry_Indicators* and *Year_Indicators* to control for potential differences in loan features across industries and over years.

IV. SAMPLE AND DESCRIPTIVE STATISTICS

Sample and Data Sources

Our initial sample consists of all public companies that have bank loan data in the Loan Pricing Corporation DealScan database for the five-year period 2005–2009. The DealScan loan data are compiled for each transaction or deal. Each deal, which is a loan contract between a borrower and bank(s) at a specific date, can have only one facility or a package of several facilities with different price and nonprice terms.[15] We consider each facility as a separate observation for our sample, as many loan characteristics and loan spreads vary across facilities. We then require that borrowers' SOX 404 disclosures are available from the Audit Analytics database before loans are initiated. The SOX 404 disclosure file from Audit Analytics provides information on the identities of disclosing companies, overall internal control effectiveness, filing dates, types and reasons of deficiencies, and so forth. We obtain borrowers' financial statement data from Compustat. Bank loan data are merged with the most recently disclosed internal control and financial statement data for the fiscal year before loans are initiated. The procedure ensures that the *Weak* measure reflects an observable result of SOX 404 compliance. We exclude financial companies from our sample and require that all loan facilities in our sample be senior debts. With regard to the types of loans, our sample includes term loans, revolvers, and 364-day facilities, but excludes bridge loans and non-fund-based facilities, such as leases and standby letters of credit. Our final sample includes 3,164 facility-years for 1,363 firms over the 2005–2009 period. Table 1 presents the sample selection procedure and the distribution of loan facilities in our sample by year and loan type. As shown in Table 1, Panel B, about 67.70 percent of loan facilities in our sample are revolvers, while about 28.03 percent and 4.27 percent are term loans and 364-day facilities, respectively.

Descriptive Statistics

Table 2, Panel A, presents descriptive statistics for all loan-specific variables at the facility level, while Panel B reports descriptive statistics for our test variables and all borrower-specific variables considered in this study. As shown in Table 2, Panel A, the mean and median of the drawn all-in spread over the LIBOR (i.e., *AIS*) are around 186 and 150 bps, respectively, with a standard deviation of about 159 bps. The mean (median) maturity is about 53 (60) months, while the mean (median) facility size is $478 million ($200 million). On average, 59.5 percent of the loan facilities in our sample have a performance pricing provision, while 68.1 percent of them have collateral. The mean and median numbers of restrictive covenants included in each loan contract are about 1.65 and 2.00, respectively, for financial covenants, and about 3.73 and 3.00, respectively, for general covenants. Most of the loan facilities in our sample are syndicated loans involving, on average, nine lenders.

As shown in Table 2, Panel B, about 9.6 percent of borrowers in the sample have a material weakness in internal control, as concluded by the auditor report under SOX 404. About 5.4 percent (4.2 percent) of facility-years in our sample have company-level (account-level) weaknesses. The *Size* variable is reasonably distributed, with a mean and median of 7.32 and 7.25, respectively. The

---

[15] For instance, a deal can comprise a line of credit facility and a term loan with different interest rates.

TABLE 1

Sample Selection and Distribution by Year and Loan Type

Panel A: Sample Selection

	Firms	Loan Facilities
Loans to public companies available in DealScan from 2005 to 2009	3,153	9,827
Less:		
Loans initiated before borrowers have SOX 404 disclosure data available in Audit Analytics	(1,069)	(3,970)
Loans borrowed by companies in the financial industry	(320)	(961)
Non-senior debts, bridge loans, bonds, letter of credit, and other non-fund-based facilities	(13)	(292)
Observations missing necessary data items for tests	(388)	(1,440)
Total observations	1,363	3,164

Panel B: Sample Distribution by Year and Loan Type

Year	Term Loans	Revolvers	364-Day Facilities	All Facilities
2005	178	461	27	666
2006	230	627	20	877
2007	264	556	35	855
2008	138	297	37	472
2009	77	201	16	294
Total	887	2,142	135	3,164
Percent (%)	28.03	67.70	4.27	100.00

mean (median) market-to-book ratio is 1.93 (1.53). On average, long-term debt, EBITDA, and tangible assets (i.e., PP&E) are about 23 percent, 13 percent, and 31 percent of total assets, respectively.

Univariate Comparisons

We first partition the full sample into two subsamples: (1) borrowers with ICW problems (n = 304), and (2) borrowers without ICW problems (n = 2,860). We then compare the differences in loan features and borrower characteristics between the two subsamples. As shown in Panel A of Table 3, the mean and median of the drawn all-in spread (*AIS*) are about 252 and 200 bps, respectively, for ICW borrowers, while they are about 179 and 150 bps, respectively, for non-ICW borrowers. Both the mean and median differences are significant at the 1 percent level, suggesting that lenders charge significantly higher loan rates to ICW borrowers than to non-ICW borrowers. With respect to other loan features, we find that, compared with loans to non-ICW borrowers, loans to ICW borrowers are smaller, have more covenants, and are more likely to be secured by collateral. We also find that ICW borrowers attract significantly fewer lenders (*NLenders*) than do non-ICW borrowers. Regarding borrower-specific characteristics, we find that, compared with borrowers without ICW problems, borrowers with ICW problems are smaller and younger, have lower growth potential, lower profitability, lower accrual quality, and higher likelihoods of default and loss, and are more likely to have foreign sales and undergoing organizational changes.

To further examine whether the nature or severity of ICW matters, we partition the ICW sample (n = 304) into two subsamples: a sample of ICW borrowers with company-level weakness

TABLE 2
Descriptive Statistics

Panel A: Loan Facility Characteristics

Variables	n	Mean	1st Quartile	Median	3rd Quartile	Std. Deviation
AIS (bps)	3,164	186.414	75.000	150.000	250.00	159.147
Maturity (months)	3,164	52.898	37.000	60.000	60.000	20.604
Loan_Size (millions)	3,164	478.473	85.000	200.000	500.000	949.409
Performance_Pricing	3,164	0.595	0.000	1.000	1.000	0.491
DSecured	2,522	0.681	0.000	1.000	1.000	0.466
FinCovIndex	3,164	1.654	0.000	2.000	3.000	1.294
GenCovIndex	3,164	3.732	1.000	3.000	6.000	2.967
NLenders	3,164	9.067	4.000	7.000	12.000	7.980

Panel B: Borrowing Firm Characteristics

Variables	n	Mean	1st Quartile	Median	3rd Quartile	Std. Deviation
Weak	3,164	0.096	0.000	0.000	0.000	0.295
Account_ICW	3,164	0.042	0.000	0.000	0.000	0.201
Company_ICW	3,164	0.054	0.000	0.000	0.000	0.226
Size	3,164	7.320	6.198	7.247	8.330	1.532
Leverage	3,164	0.230	0.073	0.200	0.331	0.203
MB	3,164	1.926	1.217	1.525	2.104	3.494
Profitability	3,164	0.129	0.089	0.123	0.177	0.136
Tangibility	3,164	0.308	0.115	0.238	0.462	0.240
CashVol	3,164	0.044	0.014	0.022	0.035	0.436
RDefault_Risk	3,164	0.501	0.222	0.444	0.778	0.320
AbsAccr	3,164	0.293	0.028	0.076	0.306	0.448
Restate	3,164	0.101	0.000	0.000	0.000	0.301
Big4	3,164	0.924	1.000	1.000	1.000	0.265
Firm_Age	3,164	24.508	10.000	17.000	39.000	17.117
NSegment	3,164	2.838	1.000	3.000	4.000	1.821
Inventory	3,164	0.114	0.009	0.082	0.171	0.125
M&A	3,164	0.261	0.000	0.000	1.000	0.439
Foreign	3,164	0.277	0.000	0.000	1.000	0.448
Restructure	3,164	0.355	0.000	0.000	1.000	0.479
Sale_Growth	3,164	0.178	0.030	0.104	0.209	0.647
Loss	3,164	0.195	0.000	0.000	0.000	0.396

(n = 171), and a sample of ICW borrowers with account-level weakness (n = 133). As shown in Table 4, Panel B, ICW borrowers with company-level weakness pay significantly higher interest rates, as reflected in higher *AIS* values, and have a higher likelihood of pledging collateral on their loans, compared with ICW borrowers with account-level weakness.

Correlation Matrix

Table 4 reports Pearson correlation coefficients among selected loan- and borrower-specific variables. As shown in Table 4, *AIS* is positively correlated with *Weak* at the 1 percent level, suggesting that banks charge higher loan rates to ICW borrowers than to non-ICW borrowers. We also find that the correlation coefficient between *AIS* and *Company_ICW* (0.13, significant at the 1

TABLE 3

Comparisons of Loan and Firm Characteristics

Panel A: Effective Internal Control versus Ineffective Internal Control

	(1) Borrowers with Effective Internal Control			(2) Borrowers with Ineffective Internal Control			Test for Difference (2) − (1)	
Variables	**n**	**Mean**	**Median**	**n**	**Mean**	**Median**	**t**	**z**
AIS (bps)	2,860	179.492	150.000	304	251.535	200.000	7.00***	8.58***
Maturity (months)	2,860	53.094	60.000	304	51.049	60.000	−1.52	−1.95*
Loan_Size (millions)	2,860	502.657	221.650	304	250.952	150.000	−9.94***	−5.55***
Performance_Pricing	2,860	0.603	1.000	304	0.520	0.000	−2.82***	−2.82***
DSecured	2,262	0.659	1.000	260	0.873	1.000	9.34***	7.02***
FinCovIndex	2,860	1.626	2.000	304	1.918	2.000	3.75***	3.59***
GenCovIndex	2,860	3.612	3.000	304	4.865	5.000	6.64***	6.77***
NLenders	2,860	9.337	7.000	304	6.530	5.000	−7.46***	−6.26***
Size	2,860	7.361	7.296	304	6.937	6.792	−4.94***	−4.92***
Leverage	2,860	0.225	0.198	304	0.273	0.225	3.02***	1.90*
MB	2,860	1.955	1.543	304	1.645	1.365	−3.60***	−5.15***
Profitability	2,860	0.133	0.127	304	0.093	0.085	−6.03***	−9.84***
Tangibility	2,860	0.308	0.237	304	0.308	0.269	0.07	−0.06
CashVol	2,860	0.045	0.021	304	0.035	0.024	−1.16	3.73***
RDefault_Risk	2,860	0.480	0.444	304	0.694	0.778	11.39***	11.07***
AbsAccr	2,860	0.275	0.072	304	0.458	0.163	5.48***	5.65***
Restate	2,860	0.066	0.000	304	0.428	0.000	12.55***	19.90***
Big4	2,860	0.926	1.000	304	0.911	1.000	−0.84	−0.90
Age	2,860	24.846	18.000	304	21.326	14.000	−3.42***	−4.45***
NSegment	2,860	2.824	3.000	304	2.970	3.000	1.33	1.62
Inventory	2,860	0.114	0.082	304	0.116	0.086	0.32	0.13
M&A	2,850	0.262	0.000	304	0.257	0.000	−0.19	−0.19
Foreign	2,860	0.264	0.000	304	0.408	0.000	4.90***	5.34***
Restructure	2,860	0.343	0.000	304	0.470	0.000	4.44***	4.42***
Sale_Growth	2,860	0.180	0.106	304	0.152	0.072	−1.14	−2.28**
Loss	2,860	0.168	0.000	304	0.447	0.000	9.49***	11.68***

Panel B: Account-Level Control Weaknesses versus Company-Level Control Weaknesses

	(1) Borrowers with Account-Level Control Weaknesses			(2) Borrowers with Company-Level Control Weaknesses			Test for Difference (2) − (1)	
Variables	**n**	**Mean**	**Median**	**n**	**Mean**	**Median**	**t**	**z**
AIS (bps)	133	220.575	175.000	171	275.614	225.000	2.89***	2.67***
Maturity (months)	133	51.053	60.000	171	51.047	60.000	−0.00	−0.73
Loan_Size (millions)	133	259.249	175.000	171	244.498	125.000	−0.43	−1.86*
Performance_Pricing	133	0.556	1.000	171	0.491	0.000	−1.13	−1.13
DSecured	109	0.789	1.000	151	0.934	1.000	3.28***	3.45***
FinCovIndex	133	1.880	2.000	171	1.947	2.000	0.44	0.80

(continued on next page)

TABLE 3 (continued)

	(1) Borrowers with Account-Level Control Weaknesses			(2) Borrowers with Company-Level Control Weaknesses			Test for Difference (2) − (1)	
Variables	**n**	**Mean**	**Median**	**n**	**Mean**	**Median**	**t**	**z**
GenCovIndex	133	4.767	5.000	171	4.942	5.000	0.48	0.71
NLenders	133	7.075	5.000	171	6.105	5.000	−1.31	−0.68
Size	133	6.953	6.741	171	6.923	6.798	−0.18	0.01
Leverage	133	0.259	0.214	171	0.285	0.243	0.84	1.23
MB	133	1.768	1.406	171	1.548	1.331	−2.02**	−2.04**
Profitability	133	0.090	0.082	171	0.095	0.088	0.38	0.26
Tangibility	133	0.348	0.334	171	0.278	0.210	−2.57**	−2.50**
CashVol	133	0.036	0.029	171	0.035	0.022	−0.28	−0.90
RDefault_Risk	133	0.680	0.778	171	0.704	0.778	0.68	0.48
AbsAccr	133	0.425	0.122	171	0.483	0.176	0.89	0.85
Restate	133	0.519	1.000	171	0.357	0.000	−2.86***	−2.83***
Big4	133	0.925	1.000	171	0.901	1.000	−0.73	−0.73
Age	133	17.752	11.000	171	24.105	16.000	3.33***	3.89***
NSegment	133	2.752	3.000	171	3.140	3.000	1.94*	1.93*
Inventory	133	0.100	0.064	171	0.128	0.103	1.84*	2.24**
M&A	133	0.211	0.000	171	0.292	0.000	1.62	1.62
Foreign	133	0.353	0.000	171	0.450	0.000	1.71*	1.70*
Restructure	133	0.481	0.000	171	0.462	0.000	−0.33	−0.33
Sale_Growth	133	0.126	0.072	171	0.173	0.072	1.17	0.72
Loss	133	0.444	0.000	171	0.450	0.000	0.12	0.12

---

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in a two-tailed test.

percent level) is much greater than that between *AIS* and *Account_ICW* (0.04, significant at the 5 percent level). Though only indicative of the underlying relation, this finding suggests that lenders are able to differentiate between borrower types based upon the nature or severity of ICW. The *Weak* variable has negative correlations with *Size* and *Profitability*, while it has positive correlations with *Leverage* and *RDefault_Risk*. The variable *AIS* is negatively correlated with *Log_Maturity*, *Log_Loan_Size*, *Log_NLenders*, and *Performance_Pricing*, while it is positively correlated with *DSecured*, *FinCovIndex*, and *GenCovIndex*.

V. REGRESSION RESULTS

ICW and Loan Spread: Test of H1

Table 5 reports the results of our main regression in Equation (1), using *AIS* as the dependent variable. In Column 1 of Table 5, we regress *AIS* on the test variable *Weak* with the full set of control variables discussed in Section III. All reported t-values are based on standard errors clustered at the firm level (Petersen 2009; Gow et al. 2010). As reported in Column 1 of Table 5, the coefficient on *Weak* is significantly positive at the 1 percent level, consistent with the prediction of H1. This result supports the view that internal control quality plays an important role in private debt contracting. Banks are unable to completely overcome information problems associated with weak internal controls, although they have superior abilities to collect and process both public and inside information about borrowers. The coefficient on *Weak* in Column 1 of Table 5 indicates that, all

TABLE 4

Pearson Correlation Matrix

Panel A: Pearson Correlation Matrix for *AIS* to *Log_Maturity*

Variables	*AIS*	*Weak*	*Account_ICW*	*Company_ICW*	*Log_Maturity*
AIS	1.00				
Weak	0.13***	1.00			
Account_ICW	0.04**	0.64***	1.00		
Company_ICW	0.13***	0.73***	−0.05***	1.00	
Log_Maturity	−0.10***	−0.02	−0.01	−0.02	1.00
Log_Loan_Size	−0.35***	−0.10***	−0.04**	−0.09***	0.16***
Log_NLenders	−0.31***	−0.10***	−0.06***	−0.08***	0.28***
Performance_Pricing	−0.28***	−0.05***	−0.02	−0.05***	0.09***
DSecured	0.47***	0.14***	0.05**	0.14***	0.10***
FinCovIndex	0.05***	0.07***	0.04**	0.05***	0.08***
GenCovIndex	0.22***	0.12***	0.07***	0.10***	0.14***
Size	−0.30***	−0.08***	−0.05***	−0.06***	−0.02
Leverage	0.17***	0.07***	0.03*	0.07***	0.05***
MB	0.02	−0.03	−0.01	−0.03	−0.01
Profitability	−0.24***	−0.09***	−0.06***	−0.06***	0.08***
Tangibility	−0.00	0.00	0.04**	−0.03*	−0.06***
CashVol	0.08***	−0.01	−0.00	−0.01	−0.00
RDefault_Risk	0.30***	0.20***	0.12***	0.15***	−0.05***
AbsAccr	0.07***	0.12***	0.06***	0.10***	−0.04**
Restate	0.04**	0.35***	0.29***	0.20***	−0.00

Panel B: Pearson Correlation Matrix for *Log_Loan_Size* to *FinCovIndex*

Variables	*Log_Loan_Size*	*Log_NLenders*	*Performance_Pricing*	*DSecured*	*FinCovIndex*
Log_Loan_Size	1.00				
Log_NLenders	0.61***	1.00			
Performance_Pricing	0.12***	0.26***	1.00		
DSecured	−0.31***	−0.24***	−0.27***	1.00	
FinCovIndex	−0.18***	0.00	0.47***	0.11***	1.00
GenCovIndex	−0.10***	0.05***	0.37***	0.41***	0.62***
Size	0.75***	0.52***	0.02	−0.38***	−0.29***
Leverage	0.12***	0.11***	−0.12***	0.17***	−0.04**
MB	−0.05**	−0.02	−0.03	−0.00	0.01
Profitability	0.17***	0.14***	0.14***	−0.14***	0.04**
Tangibility	0.17***	0.07***	0.01	−0.07***	−0.07***
CashVol	−0.06**	−0.02	−0.04**	0.03*	0.01
RDefault_Risk	−0.04**	−0.06***	−0.17***	0.23***	−0.05***
AbsAccr	−0.11***	−0.09***	−0.05***	0.09***	0.04**
Restate	−0.01	−0.04**	−0.01	0.05**	0.01

Panel C: Pearson Correlation Matrix for *GenCovIndex* to *Profitability*

Variables	*GenCovIndex*	*Size*	*Leverage*	*MB*	*Profitability*
GenCovIndex	1.00				
Size	−0.20***	1.00			
Leverage	0.06***	0.19***	1.00		

(continued on next page)

TABLE 4 (continued)

Variables	*GenCovIndex*	*Size*	*Leverage*	*MB*	*Profitability*
MB	0.02	−0.18***	−0.09***	1.00	
Profitability	−0.02	0.16***	−0.01	−0.61***	1.00
Tangibility	−0.08***	0.20***	0.27***	−0.08***	0.11***
CashVol	0.03*	−0.16***	−0.05**	0.96***	−0.70***
RDefault_Risk	0.07***	0.10***	0.62***	−0.06***	−0.37***
AbsAccr	0.05***	−0.10***	0.02	−0.00	−0.01
Restate	0.03	0.00	0.05***	−0.02	−0.05***

Panel C: Pearson Correlation Matrix for *Tangibility* to *Restate*

Variables	*Tangibility*	*CashVol*	*RDefault_Risk*	*AbsAccr*	*Restate*
Tangibility	1.00				
CashVol	−0.05***	1.00			
RDefault_Risk	0.24***	0.05***	1.00		
AbsAccr	−0.10***	−0.02	0.04**	1.00	
Restate	0.06***	−0.01	0.11***	0.01	1.00

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in a two-tailed test.

else equal, the loan spread of ICW borrowers is, on average, about 28 bps higher than that of non-ICW borrowers, which is both statistically and economically significant.[16]

Turning to control variables, we find that loan spread is negatively associated with loan facility size, the presence of performance pricing provisions, borrower size, market-to-book ratio, and profitability. Also, loan spread is positively associated with leverage, cash flow volatility, default risk, losses, term spread, and credit spread. Finally, we find that the coefficients on *Restate* and *AbsAccr* (accrual quality) are insignificant, albeit with expected positive signs.[17] All results regarding control variables are largely consistent with those of previous studies on loan pricing.

Company- versus Account-Level Material Weakness: Test of H2

To test whether banks take into account the nature or severity of material ICW when setting loan spreads, we estimate Equation (1) after replacing *Weak* by a company-level weakness indicator (*Company_ICW*) and an account-level weakness indicator (*Account_ICW*). As shown in Column 2 of Table 5, the coefficient on *Company_ICW* is significantly positive at the 1 percent level (coefficient = 43.656, t = 3.07), and the coefficient on *Account_ICW* is also positive, but insignificant (coefficient = 7.093, t = 0.58). This suggests that ICW borrowers with company-level control problems bear higher interest rates than those with account-level control problems, which is

---

[16] As reported in Table 2, the average amount of a loan facility is about $478 million for our sample, and the mean maturity is about 53 months, or 4.42 years. This means that, on average, a typical borrower with ICW has to pay higher interest expenses, of about $1.35 million per year, than a borrower without ICW over the maturity period of 4.42 years.

[17] This result is robust to different measures of accrual quality. Bharath et al. (2008) find that accrual quality is related to loan contracting terms for a sample of firms over the period 1988–2003. However, it is possible that the SOX of 2002 changed the relation between accrual quality and the cost of bank loans significantly. Note that Hutton et al. (2009) find that accrual quality is related to downside risk only before SOX, and not after. We suggest further research on the moderating effect of SOX on the relation between accrual quality and cost of debt.

TABLE 5

Relations between Loan Spreads and ICW Measures

	Model	
Variables	(1)	(2)
Test Variables		
Weak	28.189***	
	(2.72)	
Account_ICW		7.093
		(0.58)
Company_ICW		43.656***
		(3.07)
Loan-Specific Characteristics		
Log_Maturity	−10.927	−10.983
	(−1.28)	(−1.28)
Log_Loan_Size	−10.088***	−9.859***
	(−3.59)	(−3.51)
Log_NLenders	−2.281	−2.364
	(−0.59)	(−0.61)
Performance_Pricing	−30.237***	−29.994***
	(−5.84)	(−5.81)
Borrower-Specific Characteristics		
Size	−11.479***	−11.581***
	(−3.88)	(−3.89)
Leverage	55.646**	54.735**
	(2.20)	(2.16)
MB	−14.017***	−13.843***
	(−4.67)	(−4.65)
Profitability	−89.181*	−91.248*
	(−1.76)	(−1.80)
Tangibility	12.293	15.492
	(0.63)	(0.79)
CashVol	104.465***	102.778***
	(3.68)	(3.64)
RDefault_Risk	42.111**	41.597**
	(2.54)	(2.53)
AbsAccr	5.584	4.862
	(0.77)	(0.66)
Restate	4.352	5.637
	(0.53)	(0.70)
Big4	−13.197	−12.905
	(−1.00)	(−0.98)
Log_Age	−14.194***	−14.834***
	(−3.11)	(−3.27)
Log_Segment	0.661	0.512
	(0.17)	(0.13)
Inventory	−31.201	−32.231
	(−1.04)	(−1.08)
M&A	0.203	−0.100
	(0.04)	(−0.02)

(continued on next page)

TABLE 5 (continued)

Variables	Model (1)	Model (2)
Foreign	4.635	4.389
	(0.72)	(0.68)
Restructure	0.312	0.604
	(0.05)	(0.10)
Sale_Growth	−4.349*	−4.441*
	(−1.77)	(−1.81)
Loss	64.673***	64.723***
	(7.18)	(7.18)
Macroeconomic Factors		
Term_Spread	56.368***	56.762***
	(5.02)	(5.05)
Credit_Spread	32.406***	32.471***
	(2.97)	(2.98)
Intercept and Indicators		
Intercept	469.188***	468.601***
	(8.59)	(8.56)
Loan Type Indicators	Included	Included
Loan Purpose Indicators	Included	Included
Year Indicators	Included	Included
Industry Indicators	Included	Included
n	3,164	3,164
Adj. R^2	0.56	0.56

---

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in a two-tailed test.
n denotes the number of observations used in each model. The t-statistics in parentheses are based on standard errors corrected for heteroscedasticity and firm-level clustering.

consistent with H2. To directly test H2, the partial F-test shows that the coefficient on *Company_ICW* is significantly larger than that on *Account_ICW* at the 5 percent level (F = 4.34). We interpret the above results as evidence that banks are able to overcome information problems associated with account-level ICW, but not those arising from company-level ICW.

ICW and Nonprice Terms: Tests of H3 and H4

To assess the impact of ICW on the likelihood of a loan being secured by collateral, we estimate Equation (1) with *DSecured* as the dependent variable, using the probit regression procedure. Column 1a of Table 6 presents the results of this probit estimation using *Weak* as the test variable, while Column 1b reports the same using *Account_ICW* and *Company_ICW* in lieu of *Weak*. In both columns, we control for loan- and borrower-specific characteristics, determinants of ICW, and economy-wide factors. With respect to loan-specific characteristics, we include *Loan_Concentration* (the dollar amount of a loan deal divided by total assets) in lieu of *Log_Loan_Size* (the natural log of the dollar amount of the loan facility), following prior loan literature (e.g., Bharath et al. 2009). Also, we include *Prior_Lead* as an additional control variable to control for prior relationships between lead banks and borrowers. Here, *Prior_Lead* is an indicator variable that equals 1 if the lead arranger was a lead arranger for the same borrower in previous loan deals during the past five years, and 0 otherwise. Previous research suggests that banks are more likely to require collateral from borrowers that rely heavily on loans, and are less likely to require collateral from borrowers with which they have a prior banking relationship (e.g.,

TABLE 6

Effects of ICW on Collateral Requirements, Covenant Restrictions, and the Number of Lenders

Panel A: Effects of ICW for *DSecured* to *FinCovIndex*

	DSecured		*FinCovIndex*	
Variable	**(1a)**	**(1b)**	**(2a)**	**(2b)**
Test Variables				
Weak	0.413**		0.119**	
	(2.46)		(2.03)	
Account_ICW		−0.109		0.059
		(−0.60)		(0.68)
Company_ICW		0.937***		0.163**
		(3.50)		(2.29)
Loan-Specific Characteristics				
Log_Maturity	0.105	0.092	−0.015	−0.016
	(1.16)	(1.03)	(−0.38)	(−0.41)
Loan_Concentration	2.009***	2.036***		
	(4.97)	(5.02)		
Log_Deal_Size			0.027	0.028
			(0.97)	(0.99)
Log_NLenders	−0.065	−0.060	0.038	0.038
	(−0.88)	(−0.81)	(1.29)	(1.28)
Performance_Pricing	−0.420***	−0.415***	0.827***	0.829***
	(−3.75)	(−3.66)	(17.40)	(17.37)
Prior_Lead	−0.175**	−0.177**	0.015	0.016
	(−2.09)	(−2.09)	(0.47)	(0.50)
Lead_Reputation				
Borrower-Specific Characteristics				
Size	−0.323***	−0.324***	−0.171***	−0.171***
	(−5.10)	(−5.17)	(−8.06)	(−8.08)
Leverage	2.306***	2.274***	0.205	0.201
	(5.00)	(4.96)	(1.63)	(1.60)
MB	−0.258***	−0.252***	−0.098***	−0.097***
	(−4.50)	(−4.45)	(−4.98)	(−4.95)
Profitability	−1.486*	−1.524*	0.639***	0.632***
	(−1.78)	(−1.83)	(2.83)	(2.79)
Tangibility	−0.692*	−0.691*	0.026	0.032
	(−1.74)	(−1.74)	(0.22)	(0.27)
CashVol	6.425**	6.270**	0.856***	0.849***
	(2.40)	(2.33)	(5.07)	(5.04)
RDefault_Risk	1.222***	1.208***	−0.028	−0.03
	(3.92)	(3.91)	(−0.30)	(−0.32)
AbsAccr	0.164	0.172	−0.061	−0.063
	(1.41)	(1.47)	(−1.36)	(−1.40)
Restate	0.173	0.221	−0.015	−0.01
	(1.12)	(1.43)	(−0.26)	(−0.18)
Rated				

(continued on next page)

TABLE 6 (continued)

	DSecured		*FinCovIndex*	
Variable	**(1a)**	**(1b)**	**(2a)**	**(2b)**
Big4	0.189	0.220	−0.043	−0.042
	(0.96)	(1.12)	(−0.78)	(−0.77)
Log_Age	−0.383***	−0.406***	−0.049*	−0.051*
	(−5.23)	(−5.54)	(−1.69)	(−1.76)
Log_Segment	0.040	0.039	0.012	0.011
	(0.51)	(0.49)	(0.45)	(0.42)
Inventory	0.361	0.279	−0.256	−0.26
	(0.61)	(0.46)	(−1.25)	(−1.28)
M&A	0.019	0.013	0.011	0.011
	(0.19)	(0.13)	(0.32)	(0.31)
Foreign	0.012	0.008	0.015	0.014
	(0.11)	(0.07)	(0.37)	(0.36)
Restructure	−0.003	−0.013	0.006	0.007
	(−0.03)	(−0.12)	(0.17)	(0.19)
Sale_Growth	0.578***	0.557***	0.019	0.018
	(3.05)	(2.93)	(1.03)	(1.00)
Loss	0.617***	0.646***	0.058	0.059
	(3.46)	(3.59)	(1.02)	(1.03)
Macroeconomic Factors				
Term_Spread	0.302**	0.316**	0.053	0.055
	(2.01)	(2.07)	(0.96)	(0.98)
Credit_Spread	0.060	0.066	0.094**	0.095**
	(0.42)	(0.46)	(2.00)	(2.02)
Intercept and Indicators				
Intercept	8.295***	8.251***	1.012**	1.014**
	(9.45)	(9.58)	(2.08)	(2.08)
Loan Type Indicators	Included	Included	Included	Included
Loan Purpose Indicators	Included	Included	Included	Included
Year Indicators	Included	Included	Included	Included
Industry Indicators	Included	Included	Included	Included
n	2,522	2,522	3,164	3,164
Pseudo R^2	0.46	0.46	0.14	0.14

Panel B: Effects of ICW for *GenCovIndex* to *NLenders*

	GenCovIndex		*NLenders*	
Variable	**(3a)**	**(3b)**	**(4a)**	**(4b)**
Test Variables				
Weak	0.217***		−0.124**	
	(3.96)		(−1.98)	
Account_ICW		0.159*		−0.121
		(1.79)		(−1.05)
Company_ICW		0.261***		−0.126**
		(4.25)		(−2.02)
Loan-Specific Characteristics				
Log_Maturity	0.010	0.009	0.235***	0.235***
	(0.24)	(0.22)	(7.55)	(7.56)

(continued on next page)

TABLE 6 (continued)

	GenCovIndex		***NLenders***	
Variable	**(3a)**	**(3b)**	**(4a)**	**(4b)**
Loan_Concentration				
Log_Deal_Size	0.177***	0.178***	0.314***	0.314***
	(6.21)	(6.23)	(12.76)	(12.76)
Log_NLenders	−0.003	−0.004		
	(−0.11)	(−0.13)		
Performance_Pricing	0.781***	0.782***		
	(16.17)	(16.08)		
Prior_Lead	−0.007	−0.006	0.166***	0.165***
	(−0.21)	(−0.17)	(4.72)	(4.74)
Lead_Reputation			0.569***	0.569***
			(8.74)	(8.76)
Borrower-Specific Characteristics				
Size	−0.191***	−0.191***	0.047*	0.047*
	(−8.00)	(−8.01)	(1.92)	(1.92)
Leverage	0.319**	0.314**	0.221	0.221
	(2.46)	(2.42)	(1.31)	(1.31)
MB	−0.064***	−0.063***	−0.029	−0.029
	(−3.09)	(−3.06)	(−1.28)	(−1.28)
Profitability	−0.106	−0.113	−0.500**	−0.500**
	(−0.48)	(−0.51)	(−2.03)	(−2.02)
Tangibility	−0.005	0.002	0.289*	0.288*
	(−0.04)	(0.01)	(1.89)	(1.90)
CashVol	0.433**	0.424**	0.253	0.253
	(2.45)	(2.42)	(1.30)	(1.31)
RDefault_Risk	0.149	0.148	−0.15	−0.15
	(1.55)	(1.54)	(−1.36)	(−1.36)
AbsAccr	−0.039	−0.041	−0.051	−0.051
	(−0.83)	(−0.87)	(−0.90)	(−0.90)
Restate	−0.006	−0.002	−0.066	−0.066
	(−0.11)	(−0.04)	(−1.25)	(−1.27)
Rated			0.055	0.055
			(1.19)	(1.19)
Big4	0.048	0.049	0.151**	0.151**
	(0.73)	(0.75)	(2.34)	(2.35)
Log_Age	−0.062**	−0.065**	−0.03	−0.03
	(−2.09)	(−2.19)	(−1.09)	(−1.09)
Log_Segment	0.006	0.005	0.007	0.007
	(0.22)	(0.19)	(0.25)	(0.25)
Inventory	−0.081	−0.083	0.479**	0.480**
	(−0.38)	(−0.40)	(2.29)	(2.29)
M&A	0.018	0.018	0.103**	0.103**
	(0.50)	(0.50)	(2.36)	(2.36)
Foreign	0.045	0.045	0.007	0.007
	(1.09)	(1.08)	(0.15)	(0.15)
Restructure	−0.02	−0.018	−0.015	−0.015
	(−0.50)	(−0.47)	(−0.37)	(−0.37)

(continued on next page)

TABLE 6 (continued)

Variable	*GenCovIndex* (3a)	*GenCovIndex* (3b)	*NLenders* (4a)	*NLenders* (4b)
Sale_Growth	−0.008	−0.008	−0.032	−0.032
	(−0.39)	(−0.40)	(−1.12)	(−1.12)
Loss	0.186***	0.187***	−0.262***	−0.262***
	(3.42)	(3.45)	(−4.17)	(−4.20)
Macroeconomic Factors				
Term_Spread	0.067	0.069	0.035	0.035
	(1.13)	(1.16)	(0.58)	(0.58)
Credit_Spread	0.156***	0.157***	−0.054	−0.054
	(3.06)	(3.07)	(−1.35)	(−1.35)
Intercept and Indicators				
Intercept	−1.135**	−1.133**	−6.194***	−6.194***
	(−2.32)	(−2.32)	(−13.11)	(−13.11)
Loan Type Indicators	Included	Included	Included	Included
Loan Purpose Indicators	Included	Included	Included	Included
Year Indicators	Included	Included	Included	Included
Industry Indicators	Included	Included	Included	Included
n	3,164	3,164	3,164	3,164
Pseudo R^2	0.19	0.19	0.35	0.35

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in a two-tailed test.
n denotes the number of observations used in each model. Columns (1a) and (1b) are Probit regressions, and Columns (2a)–(4b) are Poisson regressions. The z-statistics in parentheses are based on standard errors corrected for heteroscedasticity and firm-level clustering.

Bradley and Roberts 2004; Vasvari 2008; Bharath et al. 2009). We therefore expect a positive (negative) sign on *Loan_Concentration* (*Prior_Lead*).

As shown in Column 1a of Table 6, the coefficient on *Weak* is significantly positive at the 5 percent level, consistent with H3. This finding suggests that lenders are more likely to require collateral for borrowers with ICW problems than for those without such problems. When *Weak* is replaced by *Account_ICW* and *Company_ICW* (as shown in Column 1b of Table 6), however, the coefficient on *Company_ICW* is significantly positive at the 1 percent level, but the coefficient on *Account_ICW* is insignificant. This suggests that the positive impact of ICW on the likelihood of a loan being secured by collateral is mainly driven by company-level weakness.

To assess the impact of ICW on the intensity or prevalence of restrictive covenants, we manually count the number of financial and general covenants included in each loan deal. We find from the DealScan database that there are a total of 30 different covenants, 18 of which are financial covenants and 12 of which are general covenants related to restrictions on prepayment, dividend, voting rights, and other restrictions.[18] We construct two covenant indices, *FinCovIndex* and *GenCovIndex*, based on counts of financial and general covenants, respectively. We then estimate Equation (1) with the two covenant indices as the dependent variables. In the covenant model, we include the size of the loan deal (*Log_Deal_Size*) in lieu of the size of the loan facility (*Log_Loan_Size*), because loan covenants are imposed at the deal level rather than at the facility level.

Columns 2a and 3a of Table 6 report the results of Poisson regressions of *FinCovIndex* and *GenCovIndex*, respectively, on *Weak* and all other control variables, while Columns 2b and 3b report the same using *Account_ICW* and *Company_ICW* in place of *Weak*. The results show that the coefficients on *Weak* are significantly positive at the 5 percent (1 percent) level in Column 2a (3a) of Table 6, consistent with H4. Columns 2b and 3b of Table 6 show that the coefficient on *Account_ICW* is significant at the 10 percent level for the general covenant regression, but insignificant for the financial covenant regression. On the other hand, the coefficients on *Company_ICW* are significantly positive at the 5 percent and 1 percent level, respectively, in Columns 2b and 3b of Table 6.

---

[18] See Bradley and Roberts (2004) for a detailed discussion on a variety of covenant restrictions used in loan contracts.

With respect to control variables, the following findings are noteworthy. First, lenders are more likely to require collateral and use general covenants for large loans. Second, collateral requirements and/or restrictive covenants are less likely to be imposed on loans to large firms, while they are more intensively used for loans to highly levered firms, younger firms, loss firms, and firms with high cash flow volatility. Third, the presence of performance pricing provisions reduces the likelihood of a loan being secured by collateral, while it increases the intensity of loan covenants. Finally, prior lead bank-borrower relationships (*Prior_Lead*) have a negative effect on the likelihood of loan collateralization, but an insignificant effect on the occurrence of restrictive covenants.

Our cross-sectional results regarding the impact of ICW on financial covenants appear to be inconsistent with those of Costello and Wittenberg-Moerman (2011). Costello and Wittenberg-Moerman (2011) focus only on ICW firms, finding that banks decrease the use of financial covenants after these firms' disclosure of ICW. In contrast, our cross-sectional tests, involving both ICW and non-ICW borrowers, show that lenders impose more financial covenants on ICW firms than on non-ICW firms.[19] Costello and Wittenberg-Moerman (2011) claim that because loans to ICW firms have a greater number of financial covenants than loans to non-ICW firms even before ICW disclosure, our findings could be due to "differences in more fundamental firm characteristics, such as riskiness and information opacity." However, we argue that this is unlikely to be the case for at least two reasons. First, in our regression model, we have carefully controlled for proxies of fundamental risk (e.g., *RDefault_Risk*), information opacity (e.g., *AbsAccr*), and many other firm characteristics. Second, we argue that even without SOX 404 (or 302) ICW disclosures *per se*, ICW could still be priced by banks, so long as banks are able to detect such weaknesses even before SOX-mandated ICW disclosures. Thus, even if Costello and Wittenberg-Moerman (2011) are correct that banks impose more financial covenants on ICW firms than on non-ICW firms before the SOX-mandated disclosure, the reason could still be weaknesses in internal controls rather than "riskiness and information opacity." Put differently, SOX 404 and 302 are not the only sources of information that banks have about borrowers' internal control effectiveness.[20]

ICW and the Number of Lenders: Test of H5

To test H5, we estimate Equation (1) using as the dependent variable the number of lenders (*NLenders*) involved in each loan deal to which a loan facility pertains. Following the prior literature (e.g., Sufi 2007; Ball et al. 2008), we further control for the presence of a prior lender-borrower relationship (*Prior_Lead*), the lead arranger's reputation (*Lead_Reputation*), and the availability of an alternative information source (*Rated*), in addition to other loan- and borrower-specific characteristics considered in the loan pricing model. Here, *Lead_Reputation* is an indicator variable that equals 1 if at least one of the lead arrangers for the loan deal was a top-25 U.S. lead arranger (in terms of loan volume) in the year immediately before the initiation of the loan, based on loan data from DealScan, and 0 otherwise. The variable *Rated* is an indicator variable that equals 1 if the borrower has a Standard & Poor's (S&P) Domestic Long Term Issuer Credit Rating, and 0 otherwise. We expect more lenders to be attracted to loans arranged by high-reputation banks and to loans to borrowers with external credit ratings by independent agencies.

---

[19] One can argue that banks should impose fewer financial covenants on ICW firms, considering such firms' lower reporting quality. We posit two explanations for the positive association between ICW and the number of financial covenants. First, by imposing more financial covenants, banks have little to lose. If accounting noise causes covenant violations, banks may be better off, because they can exercise their option to take over control rights or charge a higher loan rate via renegotiation; if the covenants are not violated due to accounting noise, these covenants are, at most, redundant. Second, banks can extract monopoly information rents based on private information they generate about the borrower (Rajan 1992). Bharath et al. (2008) argue that banks' information rent extraction is likely to be greater when accounting quality is poor, since poor accounting makes the firm more opaque to other outside stakeholders. Therefore, by imposing more financial covenants, banks can have more justifications and opportunities to obtain inside information from the borrower and extract information rents.

[20] In Section VI, we also conduct within-firm tests and further compare our study to Costello and Wittenberg-Moerman (2011).

Column 4a (4b) of Table 6 reports the results of the Poisson regression of *NLenders* on *Weak* (*Account_ICW* and *Company_ICW*) and all other control variables. As shown in Column 4a of Table 6, the coefficient on *Weak* is significantly negative at the 5 percent level, consistent with H5. When *Account_ICW* and *Company_ICW* are included in lieu of *Weak*, however, the coefficient on *Account_ICW* is insignificant, while the coefficient on *Company_ICW* is significantly negative at the 5 percent level. The above findings suggest that lenders take into account the nature or severity of material weakness in internal controls when structuring loans. The signs of the coefficients on control variables are largely consistent with prior literature.

VI. FURTHER ANALYSIS

Does SOX 404 ICW Disclosure Contain New Information for Lenders?

Our cross-sectional tests show that loan features (price terms, nonprice terms, and the number of lenders) differ significantly between ICW and non-ICW firms. A related but distinct issue is whether SOX 404 internal control disclosure contains new information for concentrated lenders with information advantage, such as banks. To examine this issue, we compare the features of a firm's loans initiated immediately before and immediately after SOX 404 disclosure. If SOX 404 disclosure provides banks with new information about borrowers' information risk, we expect to observe significant differences in loan features for the same firm before and after ICW disclosures under SOX 404.

In constructing our sample for the within-firm tests, we use each firm's *first-time* SOX 404 disclosure, because the first-time disclosure is likely to provide the most powerful setting to observe the effect of SOX 404 disclosures on loan terms. To be included in this sample, a firm must have its loans initiated in the two years before its first-time SOX 404 disclosure (the pre-SOX 404 period) and the two years after the same disclosure (the post-SOX 404 period) and these loans must have similar loan characteristics in terms of type, maturity, and amount. We match the loans initiated in the pre-SOX 404 period with those initiated in the post-SOX 404 period as closely as possible according to the following three loan characteristics: loan type (term loan or revolver), loan maturity (with a difference of less than 12 months), and loan amount (with a difference of less than 100 percent of the smaller loan). This matching procedure yields 65 ICW firms that borrow 71 pairs of loan facilities and 373 non-ICW firms that borrow 446 pairs of loan facilities. For each firm, we calculate the average values of loan features, including *AIS*, *DSecured*, *FinCovIndex*, *GenCovIndex*, and *NLenders* in the pre- and post-SOX 404 periods, and then compute their changes from the pre-SOX 404 to the post-SOX 404 period.

Table 7 presents the mean and median of these changes in loan features. Panel A of Table 7 reports the mean and median changes in five loan features for 65 ICW firms, Panel B reports the mean and median changes in five loan features for 46 ICW firms that did not previously disclose ICW problems under SOX 302, and Panel C reports the mean and median changes in five loan characteristics for 373 non-ICW firms.

TABLE 7

Changes in Loan Terms

Panel A: Changes in Loan Terms between Pre- and Post-SOX 404 Periods for 65 ICW Firms

	ΔAIS (bps)	$\Delta DSecured$	$\Delta FinCovIndex$	$\Delta GenCovIndex$	$\Delta NLenders$
Mean	22.581*	0.062	−0.831***	−0.238	−2.600
Median	0.000	0.000	0.000	0.000	0.000

Panel B: Changes in Loan Terms between Pre- and Post-SOX 404 Periods for 46 ICW Firms That Did Not Disclose ICW under SOX 302

	ΔAIS (bps)	$\Delta DSecured$	$\Delta FinCovIndex$	$\Delta GenCovIndex$	$\Delta NLenders$
Mean	34.895**	0.087*	−0.761***	−0.098	−1.543*
Median	6.250***	0.000	0.000	0.000	0.000

Panel C: Changes in Loan Terms between Pre- and Post-SOX 404 Periods for 373 Non-ICW Firms

	ΔAIS (bps)	$\Delta DSecured$	$\Delta FinCovIndex$	$\Delta GenCovIndex$	$\Delta NLenders$
Mean	−24.967***	0.017	−0.435***	−0.240*	−0.316
Median	−25.000***	0.000	0.000	0.000	0.000

Panel D: Changes in Loan Terms between Pre-Remediation and Post-Remediation Periods for 32 Remediating Firms

	ΔAIS (bps)	$\Delta DSecured$	$\Delta FinCovIndex$	$\Delta GenCovIndex$	$\Delta NLenders$
Mean	−22.078**	−0.130*	−0.656*	−1.641**	−2.375
Median	−7.500***	0.000	0.000	0.000	0.000

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent levels, respectively, in a two-tailed test.
Panel A presents the mean and median of the changes in loan characteristics for 65 ICW firms within two years before and after the firms' first-time SOX 404 disclosures. Panel B presents the mean and median of the changes in loan characteristics for 46 ICW firms that did not disclose ICW under SOX 302. Panel C presents the mean and median of the changes in loan characteristics for 373 non-ICW firms within two years before and after the firms' first-time SOX 404 disclosures. Panel D presents the mean and median of the changes in loan characteristics for 32 firms around their remediating ICW.

As shown in Table 7, Panel A, the mean *AIS* change for 65 ICW firms is nearly 23 bps and is significant at the 10 percent level. This is in line with the view that, on average, loan rates charged to ICW firms increase after ICW disclosure under SOX 404. For ICW firms that did not disclose ICW under SOX 302 before the first-time SOX 404 disclosure, as presented in Table 7, Panel B, the mean (median) *AIS* changes are about 35 (six) bps, and significant at less than the 5 percent (1 percent) level. The stronger *AIS* increase in Panel B than in Panel A of Table 7 suggests that lenders react more to ICW disclosures under SOX 404 when the disclosures were not made in the past and, thus, are more likely to convey new information.[21]

---

[21] Note that SOX 404 disclosure itself carries some confirmation effect (or feedback value), even if the firm previously disclosed ICW through unaudited disclosure. Consistent with this conjecture, we find that banks react to a firm's SOX 404 ICW disclosures even if the firm previously disclosed ICW under SOX 302.

As reported in Table 7, Panel C, both the mean and median *AIS* changes for non-ICW firms are about −25 bps. These changes are significant at the 1 percent level, suggesting that loan rates charged to non-ICW firms *decrease* after firms assure the inexistence of ICW problems by disclosing auditor-attested effectiveness of internal controls under SOX 404. The above results, taken together, suggest that SOX 404 internal control disclosures provide banks with new information about borrowers' information risk; that is, banks are not able to completely obtain all information regarding the effectiveness of a firm's internal control system before mandatory SOX 404 disclosure.[22]

With regard to the change in the number of financial covenants, Panels A and B of Table 7 show that the mean change in *FinCovIndex* is significantly negative at the 1 percent level, suggesting that lenders impose fewer financial covenants on loans to ICW firms after SOX 404 disclosures. This is consistent with the result of Costello and Wittenberg-Moerman (2011), who interpret their similar finding as evidence that lenders use fewer financial covenants for ICW firms because the accounting numbers used in loan contracts are less reliable for these firms. As shown in Panel C of Table 7, however, the change in the number of financial covenants for non-ICW firms is significantly negative, as well, at the 1 percent level. This result casts some doubt on the interpretation of Costello and Wittenberg-Moerman (2011) regarding the decrease in the use of financial covenants for ICW firms from the pre- to the post-SOX 404 period. The similar pattern of change in the number of financial covenants, that is, Δ*FinCovIndex*, for *both* ICW and non-ICW firms observed in Panels A through C of Table 7 could be related to lenders' improved knowledge about borrowers. As banks become familiar with a firm over time, it would be less beneficial for banks to *ex post* monitor the firm using financial covenants.

Moreover, the decreased use of financial covenants for both ICW and non-ICW firms could also be driven, at least partially, by a general regulatory effect of SOX. That is, SOX-mandated regulations could have improved the overall information environment and corporate governance of U.S. firms and, thus, could have reduced the need for bank monitoring through financial covenants. In addition, we find that the number of financial covenants decreases more (significant at the 10 percent level) for ICW firms than for non-ICW firms. This difference could be interpreted as evidence that the general regulatory impact of SOX on ICW firms is more pronounced because ICW firms are likely to have higher *ex ante* information asymmetry.[23,24]

In addition, Panel B of Table 7 shows that the mean change in *DSecured* is positive and weakly significant at the 10 percent level, while the mean change in *NLenders* is negative and weakly significant at the 10 percent level. This result suggests that loans to ICW firms (under SOX 404) that did not disclose ICW under SOX 302 are more likely to be secured by collateral and attract fewer lenders. In Panel C of Table 7, we also find that fewer general covenants are imposed on loans to non-ICW firms in the post-SOX 404 disclosure period than in the pre-SOX 404 disclosure period.

Impact of ICW Remediation on Loan Features

In this subsection, we conduct another within-firm analysis to examine whether loan rates charged to ICW borrowers decrease after they remediate their internal control problems. For this purpose, we identify a sample of firms that disclose ICW under SOX 404 in a year and then report that their internal controls are effective in a later year. In so doing, we further require these remediating firms to have loans initiated in both the pre-remediation period (from the disclosure date of ICW to the remediation date)[25] and the post-remediation period (from the remediation date to two years after). Following the same procedures described in the last subsection, we match the loans in the two periods and compute the mean and median changes in loan characteristics for the remediating firms.

---

[22] Again, this "new" information can also stem from the confirmation value of the audited internal control information.

[23] In the extreme case, bank monitoring through covenants will not be necessary for a firm with absolute *ex ante* transparency and, thus, SOX will not have an impact on such firms.

[24] Given these conjectures, we believe that both the cross-sectional test and the time-series test have their own merits and limitations. Thus, it seems necessary to conduct both tests to obtain a more complete picture of the relation between ICW and bank loan contracting.

The results, as reported in Table 7, Panel D, show that the mean and median changes in loan rate around the ICW remediation for 32 firms are −22.078 and −7.500, respectively, significant at the 5 percent and 1 percent levels, respectively. The significant decrease in loan rate in the post-remediation period suggests that firms enjoy a meaningful reduction in the cost of bank loans after ICW remediation. With regard to changes in nonprice terms, we find that compared with loans in the pre-remediation period, loans initiated in the post-remediation period are less likely to be secured by collateral (at the 10 percent level) and tend to include fewer financial covenants (at the 10 percent level) and fewer general covenants (at the 5 percent level). Overall, the analyses of remediating firms provide further evidence that auditor-attested internal control disclosures under SOX 404 are informative to banks.

Using a sample of SOX 302 disclosures, Costello and Wittenberg-Moerman (2011) conclude that "there is no pricing effect after the ICW is corrected," which appears to be inconsistent with our results. Note, however, that their study compares interest rates for loans initiated after ICW remediation (the "corrected" period) to those initiated before ICW disclosure (the "prior" period).[26] In contrast, we compare the interest rates for loans initiated after ICW remediation with those for loans initiated after ICW disclosure but before ICW remediation. As a matter of fact, Costello and Wittenberg-Moerman's (2011) results are somewhat consistent with ours: For example, in Column 1 of their Table 4, the coefficient on *Uncorrected* is about 29 bps and significantly positive, while the coefficient of *Corrected* is nearly zero and insignificant. Therefore, the insignificant coefficient for the *Corrected* indicator variable can also be interpreted in such a way that the loan rate decreases significantly relative to that in the uncorrected period after ICW remediation, and returns to its prior period level.

ICW and Default Risk

An important remaining challenge is to rule out the possibility that our test variable *Weak* simply captures differences in default risk (rather than the lack of reporting system quality) between ICW and non-ICW firms, thereby leading us to observe differences in various loan terms between the two. To address this possibility, we examine whether *Weak* has the ability to predict the actual occurrence of defaults observed for our sample firms. If *Weak* is merely a proxy for default risk, but not for reporting system quality, we expect it to be a significant predictor of actual defaults or bankruptcies. If *Weak* captures information risk rather than the measured default risk, we expect its predictive ability with respect to actual defaults to be insignificant. Following the recent bankruptcy prediction literature (e.g., Shumway 2001; Chava and Jarrow 2004; Chava et al. 2009), we estimate the Cox (1972) proportional hazard model to assess the ability of our ICW measures to predict actual defaults observed for the borrowing firms in our sample. We obtain data on actual default events from Moody's Corporate Default Risk Service. The results from the hazard model estimations, though not reported here, show no evidence that *Weak*, *Company_ICW*, and *Account_ICW* predict the occurrence of actual defaults. This finding is inconsistent with the view that our ICW measures simply capture default risk rather than information risk associated with the lack of adequate internal controls.

---

[25] The remediation date refers to the date of the firm's first clean-opinion internal control disclosure under SOX 404 after its previous ICW disclosure. The average length of the pre-remediation period for 32 firms in our sample is about 463 days, or 1.27 years.

[26] Costello and Wittenberg-Moerman (2011) include both the *Uncorrected* and *Corrected* indicator variables in the model, which means the loan rate in the prior period is used as the benchmark to interpret the coefficients for the *Uncorrected* and *Corrected* indicator variables. See Costello and Wittenberg-Moerman (2011, Table 4) for details.

Deal-Level Analyses

Thus far, we have conducted our empirical analyses at the loan facility level; in other words, we consider each loan facility to be an independent observation. However, facility-level loan features in a deal may not be independent, as borrowers may have negotiated loan terms with lenders at the deal level. To address this concern, we construct a reduced sample of 2,334 observations at the loan-deal level and then re-estimate our main regressions. In so doing, we compute the average of the facility-level values of various loan features across multiple facilities in a deal, using facility size as a weight. Though not reported here, the inferences from the deal-level regression results are identical to those from the facility-level regression results reported in Tables 5 and 6.

Joint Determination of Loan Terms

To some extent, various loan contract terms can be jointly determined. However, the syndicated loan literature indicates that nonprice terms are normally determined before setting the loan interest rate in the syndication process (e.g., Dennis et al. 2000; Bharath. 2009; Ivashina 2009). Nonetheless, we address the possible joint determination of price and nonprice terms by estimating a system of equations using the seemingly unrelated regression (SUR) procedure.[27] Unreported results show that the SUR results are very similar to those reported in Tables 5 and 6.[28]

VII. CONCLUSIONS

Using a sample of 3,164 loan facility-years for borrowers that filed SOX 404 disclosures with the SEC during 2005–2009, this study compares the price and nonprice terms of loan contracts between borrowers with ICW problems and borrowers without such problems, after controlling for loan-specific, borrower-specific, and economy-wide factors known to affect contract terms. We find that ICW firms incur higher direct and indirect costs of bank loans than non-ICW firms. In addition, using a within-firm analysis, we show that banks charge significantly higher interest rates for loans initiated after the disclosure of ICW under SOX 404 than for loans initiated before the disclosure. This finding suggests that SOX 404 disclosure is informative to banks and other private lenders. Finally, we find evidence that banks charge lower loan rates for ICW firms after the remediation of internal control problems reported under SOX 404.

Overall, our results indicate that banks take into account the quality of internal control over financial reporting when setting the price and nonprice terms of loan contracts. This finding suggests that credit stakeholders in the private debt market view material weaknesses in internal controls over financial reporting as an information-risk-increasing factor that is incrementally significant, beyond traditional credit risk factors.

---

[27] The SUR equations include models with the following dependent variables: loan pricing, collateral, financial covenants, general covenants, and number of lenders.

[28] As an additional robustness test for loan pricing results, we control for nonprice terms in the loan pricing model and find that the results reported in Table 5 are robust to the inclusion of nonprice terms as additional loan-specific controls (untabulated here).

REFERENCES

Altman, E. 1968. Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance* 23 (4): 589–609.

Ashbaugh-Skaife, H., D. W. Collins, and W. R. Kinney. 2007. The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics* 44 (1–2): 166–192.

Ashbaugh-Skaife, H., D. W. Collins, W. R. Kinney, and R. LaFond. 2008. The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review* 83 (1): 217–250.

Ashbaugh-Skaife, H., D. W. Collins, W. R. Kinney, and R. LaFond. 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity. *Journal of Accounting Research* 47 (1): 1–34.

Ball, R., R. M. Bushman, and F. P. Vasvari. 2008. The debt-contracting value of accounting information and loan syndicate structure. *Journal of Accounting Research* 46 (2): 247–387.

Ball, R., and L. Shivakumar. 2006. The role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research* 44 (2): 207–242.

Beneish, M. D., M. B. Billings, and L. D. Hodder. 2008. Internal control weakness and information uncertainty. *The Accounting Review* 83 (3): 665–703.

Berger, A. N., and G. F. Udell. 1990. Collateral, loan quality, and bank risk. *Journal of Monetary Economics* 25 (1): 21–42.

Bharath, S. T., S. Dahiya, A. Saunders, and A. Srinivasan. 2009. Lending relationships and loan contract terms. *Review of Financial Studies* (forthcoming).

Bharath, S. T., J. Sunder, and S. V. Sunder. 2008. Accounting quality and debt contracting. *The Accounting Review* 83 (1): 1–28.

Bradley, M., and M. R. Roberts. 2004. The structure and pricing of corporate debt covenants. Working paper, Duke University and the University of Pennsylvania.

Chava, S., and R. Jarrow. 2004. Bankruptcy prediction with industry effects. *Review of Finance* 8 (4): 537–569.

Chava, S., D. Livdan, and A. Purnanandam. 2009. Do shareholder rights affect the cost of bank loans? *Review of Financial Studies* 22 (8): 2973–3004.

Cole, R. A. 1998. The importance of relationships to the availability of credit. *Journal of Banking and Finance* 22 (6–8): 959–977.

Costello, A. M., and R. Wittenberg-Moerman. 2011. The impact of financial reporting quality on debt contracting: Evidence from internal control weakness reports. *Journal of Accounting Research* 49 (1): 97–136.

Cox, D. R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society*, Series B (Methodological) 34 (2): 187–220.

Dechow, P., W. Ge, and C. Schrand. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics* 50 (2–3): 344–401.

Dechow, P., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

Denis, D. J., and V. T. Mihov. 2003. The choice among bank debt, non-bank private debt, and public debt: Evidence from new corporate borrowings. *Journal of Financial Economics* 70 (1): 3–28.

Dennis, S., and D. J. Mullineaux. 2000. Syndicated loans. *Journal of Financial Intermediation* 9 (4): 404–426.

Dennis, S., D. Nandy, and I. G. Sharpe. 2000. The determinants of contract terms in bank revolving credit agreements. *Journal of Financial and Quantitative Analysis* 35 (1): 87–110.

Dhaliwal, D., C. Hogan, R. Trezevant, and M. Wilkins. 2011. Internal control disclosures, monitoring, and the cost of debt. *The Accounting Review* 86 (4): 1131–1156.

Doyle, J., W. Ge, and S. McVay. 2007a. Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics* 44 (1–2): 193–223.

Doyle, J., W. Ge, and S. McVay. 2007b. Accruals quality and internal control over financial reporting. *The Accounting Review* 82 (5): 1141–1170.

Duffie, D., and D. Lando. 2001. Term structures of credit spreads with incomplete accounting information. *Econometrica* 69 (3): 633–664.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* 59 (4): 1553–1583.

Feng, M., C. Li, and S. McVay. 2009. Internal control and management guidance. *Journal of Accounting and Economics* 48 (2–3): 190–209.

Fama, E. 1985. What's different about banks? *Journal of Monetary Economics* 15 (1): 29–39.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39 (2): 295–327.

Gow, I. D., G. Ormazabal, and D. J. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.

Graham, J. R., S. Li, and J. Qiu. 2008. Corporate misreporting and bank loan contracting. *Journal of Financial Economics* 89 (1): 44–61.

Hammersley, J. S., L. A. Myers, and C. Shakespeare. 2008. Market reactions to the disclosure of internal control weaknesses and to the characteristics of those weaknesses under Section 302 of the Sarbanes-Oxley Act of 2002. *Review of Accounting Studies* 13 (1): 141–165.

Hutton, A. P., A. J. Marcus, and H. Tehranian. 2009. Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics* 94 (1): 67–86.

Ivashina, V. 2009. Asymmetric information effects on loan spreads. *Journal of Financial Economics* 92 (2): 300–319.

Jimenez, G., V. Salas, and J. Saurina. 2006. Determinants of collateral. *Journal of Financial Economics* 81 (2): 255–281.

Kim, J-B., J. S. L. Tsui, and C. H. Yi. 2010. The voluntary adoption of international financial reporting standards and loan contracting around the world. *Review of Accounting Studies* (forthcoming).

Lambert, R., C. Leuz, and R. Verrecchia. 2007. Accounting information, disclosure, and cost of capital. *Journal of Accounting Research* 45 (2): 385–420.

Moody's Investors Service. 2004. *First Year Section 404 Reports on Internal Control: Impact on Ratings Will Depend on Nature of Material Weaknesses Reported.* New York, NY: Moody's Investors Service.

Moody's Investors Service. 2006. *Second Year Section 404 Reports on Internal Control: Delinquent Filers are most at Risk of Negative Rating Action.* New York, NY: Moody's Investors Service.

Moody's Investors Service. 2007. *Third Year Section 404 Reports on Internal Control: Controls Problems are Decreasing, but Reporting can be Improved.* New York, NY: Moody's Investors Service.

Ogneva, M., K. R. Subramanyam, and K. Raghunandan. 2007. Internal control weaknesses and cost of equity: Evidence from SOX Section 404 disclosures. *The Accounting Review* 82 (5): 1255–1297.

Ohlson, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18 (1): 109–131.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparison approach. *Review of Financial Studies* 22 (1): 435–480.

Public Company Accounting Oversight Board (PCAOB). 2007. *An Audit of Internal Control over Financial Reporting that is Integrated with an Audit of Financial Statements.* Auditing Standard No. 5. Washington, D.C.: PCAOB.

Rajan, R. G. 1992. Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance* 47 (4): 1367–1400.

Rajan, R. G., and A. Winston. 1995. Covenants and collateral as incentives to monitor. *Journal of Finance* 50 (4): 1113–1146.

Shumway, T. 2001. Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business* 74 (1): 101–124.

Smith, C. W., and J. B. Warner. 1979. On financial contracting: An analysis of bond covenants. *Journal of Financial Economics* 7 (2): 117–161.

Strahan, P. E. 1999. Borrower risk and the price and nonprice terms of bank loans. Working paper, Federal Reserve Bank of New York.

Sufi, A. 2007. Information asymmetry and financing arrangements: Evidence from syndicated loans. *Journal of Finance* 62 (2): 629–668.

U. S. House of Representatives. 2005. *The Impact of the Sarbanes-Oxley Act.* Hearing before the Committee on Financial Services (April 21). Washington, D.C.: Government Printing Office.

Vasvari, F. P. 2008. Equity compensation and the pricing of syndicated loans. Working paper, London Business School.

APPENDIX A

Variable Definitions

ICW Variables	
Weak	An indicator variable that equals 1 if the auditor concludes a firm's internal control over financial reporting is not effective under SOX 404, and 0 otherwise. The data on SOX 404 disclosures are from the Audit Analytics database.
Account_ICW	An indicator variable that equals 1 if a borrowing firm reports any account-level ICW, and 0 otherwise. Data on SOX disclosures are from the Audit Analytics database.
Company_ICW	An indicator variable that equals 1 if a firm reports any company-level ICW or did not complete an internal control assessment, and 0 otherwise. Data on SOX disclosures are from the Audit Analytics database.
Borrower-Specific Variables	
Size	Firm size, which is the natural log of total assets in millions of dollars (Compustat data item 6).
Leverage	Leverage ratio, defined as the long-term debt (Compustat data item 9) divided by total assets (Compustat data item 6).
MB	Market-to-book ratio, measured as the market value of equity plus the book value of debt (Compustat data item 6 − Compustat data item 60 + Compustat data item 25 × Compustat data item 199) divided by total assets (Compustat data item 6).
Profitability	EBIDTA (Compustat data item 13) divided by total assets (Compustat data item 6).
Tangibility	Net PP&E (Compustat data item 8) divided by total assets (Compustat data item 6).
CashVol	Cash flow volatility, measured by the standard deviation of quarterly cash flows from operations (change in quarterly Compustat data item 108) scaled by total assets (Compustat data item 6) over the past five fiscal years.
RDefault_Risk	Decile rank of the first principal component of three commonly used default risk proxies: Altman's (1968) z-score, Ohlson's (1980) O-score, and Shumway's (2001) probability of bankruptcy (−0.7999 × *z-score* + 0.8948 × *O-score* + 0.6239 × *Probability of Bankruptcy*). Larger values indicate higher default risk.
AbsAccr	The absolute value of abnormal accruals obtained from the modified Jones model (Dechow et al. 1995) considering accounting conservatism (Ball and Shivakumar 2006).
Restate	An indicator variable that equals 1 if the firm restates its financial statements in the year of disclosing ICWs under SOX 404.
Rated	An indicator variable that equals 1 if the borrower has an S&P Domestic Long Term Issuer Credit Rating (Compustat data item 280), and 0 otherwise.
Big4	An indicator variable that equals 1 if the auditor of a borrower is one of the Big 4 auditors, and 0 otherwise.
Age	The number of years the firm has data in Compustat.
Log_Age	The natural log of *Age*.
NSegment	The number of the firm's business segment reported by Compustat segments.
Log_Segment	The natural log of *NSegment*.
Inventory	Inventory (Compustat data item 3) divided by total assets (Compustat data item 6).
M&A	An indicator variable that equals 1 if the firm is involved in mergers or acquisitions, and 0 otherwise (Compustat AFNT item 1).

Foreign	An indicator variable that equals 1 if the firm has a nonzero foreign currency translation (Compustat data item 150), and 0 otherwise.
Restructure	An indicator variable that equals 1 if any of Compustat data items 376, 377, 378, and 379 are nonzero, and 0 otherwise.
Sale_Growth	Growth rate in sales (Compustat data item 12).
Loss	An indicator variable that equals 1 if the firm has negative aggregate earnings before extraordinary items (Compustat data item 18) in the last and current fiscal years, and 0 otherwise.
Loan-Specific Variables	
AIS	The dependent variable, which is measured by the drawn all-in spread charged by the bank over the LIBOR for the drawn portion of the loan facility, obtained from the DealScan database.
Maturity	The maturity of the loan in months.
Log_Maturity	The natural log of *Maturity*.
Loan_Size	The amount of the loan facility in millions of dollars.
Log_Loan_Size	The natural log of *Loan_Size*.
Deal_Size	The dollar amount of the loan deal.
Log_Deal_Size	The natural log of *Deal_Size*.
NLenders	The number of banks in the loan deal.
Log_NLenders	The natural log of *NLenders*.
Performance_Pricing	An indicator variable that equals 1 if the loan contract includes performance pricing provisions, and 0 otherwise.
DSecured	An indicator variable that equals 1 if the loan facility is secured with collateral, and 0 otherwise.
FinCovIndex	Financial covenant index constructed by counting the number of financial covenants included in a loan contract.
GenCovIndex	General covenant index constructed by counting the number of general covenants included in a loan contract.
Loan_Concentration	*Deal_Size* divided by the sum of *Deal_Size* and the borrower's total liabilities (Compustat data item 181).
Prior_Lead	An indicator variable that equals 1 if at least one of the lead arrangers for the current deal has been a lead arranger of previous deals for the same borrower during the past five years, and 0 otherwise.
Lead_Reputation	An indicator variable that equals 1 if at least one of the lead arrangers for the loan deal was a top-25 U.S. lead arranger (in terms of loan volume) in the year before the initiation of the loan based on the loan data from DealScan, and 0 otherwise.
Loan_Purpose_Indicators	A series of the indicator variables for the purposes of loan facilities in DealScan, including corporate purposes, debt repayment, working capital, CP backup, takeover, and acquisition line.
Loan_Type_Indicators	A series of indicator variables for the types of loan facilities in DealScan, including term loan, revolvers, and 364-day facilities.
Macroeconomic Variables	
Term_Spread	Difference in the yield between ten-year and two-year U.S. Treasury bonds measured one month before the loan becomes active, obtained from the Federal Reserve Board of Governors.
Credit_Spread	Difference in the yield between BAA- and AAA-rated corporate bonds measured one month before the loan becomes active, obtained from the Federal Reserve Board of Governors.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1189–1212

American Accounting Association
DOI: 10.2308/accr-10037

Modeling Discretionary Accrual Reversal and the Balance Sheet as an Earnings Management Constraint

William R. Baber
Georgetown University

Sok-Hyon Kang
Ying Li
The George Washington University

ABSTRACT: This study presents conceptual and empirical analyses of discretionary accrual reversal in the earnings management context. We specifically focus on the extent that income-increasing (decreasing) discretionary accruals initiated in a prior period reverse to become income-decreasing (increasing) accruals in the current period. The analysis suggests that the extent that such reversals constrain the ability to manage toward earnings objectives depends on both the magnitude of past accrual-based earnings management and the *reversal speed* of past discretionary accruals. To demonstrate the empirical implications of the analysis, we consider discretionary accrual reversal speed as an additional determinant of the balance sheet constraint on earnings management (Barton and Simko 2002). We show that, conditional on the magnitude of net operating asset overstatement, the probability of achieving quarterly earnings forecasts varies inversely with reversal speed.

Keywords: *earnings management; discretionary accruals; discretionary accrual reversal; reversal speed.*

JEL Classifications: *M4; M41.*

Data Availability: *Data are available from public sources identified in the paper.*

The authors appreciate comments by Andy Call, Doug Hanna, Weining Zhang, workshop participants at Baruch College–CUNY, George Mason University, The George Washington University, Korea University Business School, National University of Singapore, Temple University, Texas Christian University, Southern Methodist University, University of Maryland, and the 2010 American Accounting Association Annual Meeting. We thank the two anonymous referees, Sanjay Kallapur (editor), and Steven Kachelmeier (senior editor) for their thorough and constructive comments.

Editor's note: Accepted by Sanjay Kallapur.

Submitted: June 2010
Accepted: January 2011
Published Online: April 2011

I. INTRODUCTION

The reversal of accounting accruals is a well-known and frequently studied phenomenon in the accounting literature,[1] yet relatively few empirical studies explicitly model the reversal of discretionary accruals in the context of earnings management. The objectives of this study are to model the reversal process and to examine empirically how reversals of discretionary accruals initiated in prior periods influence the ability to manage future earnings.

We construct a simple characterization in which discretionary accruals created in one period reverse in their entirety n periods into the future. We designate the *speed* of discretionary accrual reversal as $1/n$, as smaller n implies a faster reversal. We then show how, due to double-entry bookkeeping under accrual accounting, earnings manipulations can accumulate on the balance sheet, creating a bank of cumulative discretionary accruals that carries forward into future periods and constrains the ability to manage future income when prior discretionary accruals eventually reverse back to the income statement. We predict that, holding constant the accumulated bank of discretionary accruals, the constraint on earnings management due to discretionary accrual reversal increases with reversal speed.

To test this prediction empirically, we extend the analysis in Barton and Simko (2002), who argue that, because managers have limited reporting discretion, optimistic reporting using discretionary accruals constrains the ability to achieve future earnings objectives. Moreover, under the double-entry bookkeeping system, the balance sheet contains cumulative past income statement biases. Taken together, the authors predict and find that the ability to achieve earnings targets through earnings management varies inversely with beginning-of-the-period net operating assets (NOA), which is a proxy for prior financial reporting bias. However, Barton and Simko (2002) do not articulate the link between NOA and the constraint on future earnings management, nor do they consider the speed with which discretionary accruals reverse. In contrast, by explicitly considering discretionary accrual reversals, we provide a link between NOA and subsequent earnings management.

Earnings are typically disclosed quarterly, such that we use 1993–2007 quarterly earnings announcements to test our prediction. Cohen et al. (2008) find that the use of accruals to manage earnings declines sharply after the Sarbanes-Oxley Act (SOX), so we partition the data into pre-SOX and post-SOX periods. We find that negative associations between NOA and the ability to meet earnings expectations documented by Barton and Simko (2002) are much less substantial in the post-SOX period. Thus, the empirical analyses of discretionary accrual reversals focus on the 1993–2003 sub-period. Consistent with expectations predicated on formal analysis, we find an inverse association between the speed of discretionary accrual reversal and the probability of meeting or beating quarterly analyst forecasts (MBE). That is, conditional on beginning NOA, the ability to MBE decreases with estimated reversal speed.

This study makes two contributions to the earnings management literature. First, to our knowledge, this is the first empirical study that explicitly models the reversal of discretionary accruals in the earnings management context. We develop empirical measures to capture the reversal speed, and provide a testable hypothesis about how reversal affects the ability to bias earnings. Second, by articulating the link between overstated NOA and subsequent earnings management, we offer a conceptual foundation for interpreting Barton and Simko's (2002) result that balance sheets constrain future earnings management.

---

[1] See Guay et al. (1996), Hunt et al. (1996), Beneish (1997), DeFond and Park (2001), Barton and Simko (2002), and Kothari et al. (2005). Theoretical studies that model the reversal process include Verrecchia (1986), Dye (1988), Trueman and Titman (1988), and Kirschenheiter and Melumad (2002). Most of these studies examine the intertemporal allocation of income (e.g., income smoothing), rather than how the reversal of prior discretionary accruals affects the ability to manipulate earnings in the current period.

Section II develops the hypothesis and lays out the framework for empirical tests. Section III outlines the research design. Section IV describes the sample, and Section V reports results. Concluding remarks are in Section VI.

II. HYPOTHESIS AND MODEL DEVELOPMENT

Accrual Reversal as an Earnings Management Constraint

Similar to previous research, (e.g., DeFond and Park 2001; Barton and Simko 2002), we consider a setting in which a firm uses accruals within the boundaries of GAAP to manage reported income over a sustained period.[2] As is well known, accrual-based earnings management (EM) under the double-entry bookkeeping system implies that the firm's balance sheet is also impacted by past earnings management (Kang and Sivaramakrishnan 1995; Beneish 1997; Barton and Simko 2002; DeFond 2002).

More formally, consider the following representation of period t reported earnings (X_t), which consists of unmanaged earnings (X_t^*) and managed accruals:

$$X_t = X_t^* + (d_t^{new} - d_t^{reversed}), \tag{1}$$

where d_t^{new} and $d_t^{reversed}$, respectively, denote discretionary accruals initiated in period t, and past discretionary accruals that reverse in period t. Together, d_t^{new} and $d_t^{reversed}$ constitute *discretionary accruals* in the extant literature (Jones 1991; Dechow et al. 1995; Kothari et al. 2005). Three points about this characterization are worth noting.

First, given the constraints imposed by double-entry accounting, the reversal of prior discretionary accruals, $d_t^{reversed}$, is an unavoidable component of reported earnings. That is, the reversal of prior EM is *not* "discretionary" in period t. In contrast with most EM studies, we distinguish discretionary accruals initiated in period t (d_t^{new}) from period t reversals of discretionary accruals initiated in prior periods ($d_t^{reversed}$). Accordingly, we designate d_t^{new} as "discretionary" accruals (DA), and ($d^{new} - d^{reversed}$) as "net" discretionary accruals (NDA).

Second, also because of double-entry bookkeeping under the accrual accounting system, the extent that earnings are managed continuously manifests as *cumulative* overstatement of NOA at the *end* of period t, $\sum_{k=0}^{t-1}(d_{t-k}^{new} - d_{t-k}^{reversed})$, denoted CDA_t (cumulative discretionary accruals). Notice that the period t *change* in the CDA balance (ΔCDA_t) represents *net* discretionary accruals ($d_t^{new} - d_t^{reversed}$)—i.e., $\Delta CDA_t = d_t^{new} - d_t^{reversed}$. As a result, the widely used approach of using primarily the *change* in the NOA balance (which includes the *change* in the CDA balance) to isolate discretionary from non-discretionary accruals (e.g., Jones 1991; Dechow et al. 1995; Kothari et al. 2005) yields an estimate of *net* discretionary accruals, not *new* discretionary accruals initiated in period t. By contrast, the balance sheet-based approach, as adopted by Kang and Sivaramakrishnan (1995), DeFond and Park (2001), and White (2010),[3] emphasizes estimating *new* discretionary accruals.[4]

---

[2] Notice that accrual-based earnings management can include "real" earnings management that affects both NOA balance and income statement, e.g., channel stuffing.

[3] For example, White (2010), which extends Francis and Smith (2005), treats reversed discretionary accruals as a measurement error in the treatment variable, i.e., as something to be removed, not to be estimated. White (2010) uses an errors-in-variable approach to purge reversed discretionary accruals. Whether to estimate d_t^{new} or $d_t^{new} - d_t^{reversed}$ is a design choice, to be tailored for the specific hypothesis in question.

[4] To clarify, assume that $\phi_t (0 \leq \phi_t \leq 1)$ of the beginning CDA balance reverses in period t (i.e., $d_t^{reversed} = \phi_t CDA_{t-1}$). Then, the ending CDA balance, CDA_t, is $d_t^{new} + (1 - \phi_t)\ CDA_{t-1}$. If $\phi_t = 1$ (i.e., the entire CDA balance reverses), or if $CDA_{t-1} = 0$ (*EM* occurs in period t only, as in Kang and Sivaramakrishnan 1995; DeFond and Park 2001; White 2010), then $CDA_t = d_t^{new}$. Otherwise, $CDA_t = d_t^{new} + (1 - \phi_t)\ CDA_{t-1} = d_t^{new} + CDA_{t-1} - d_t^{reversed}$ and, thus, the balance sheet approach may not yield a unique estimate of either d_t^{new} or $d_t^{new} - d_t^{reversed}$.

Third, past reporting bias toward one direction constrains the ability to bias current earnings in the same direction in two ways. More specifically, the first constraint arises when a fraction of CDA_{t-1} must be reversed back to the income statement in period t ($d_t^{reversed}$). This reversal, which we designate the *EM Constraint*, and which constrains period t net income dollar-for-dollar, is the focus of this study.

Barton and Simko (2002) note that, because the balance sheet reflects the extent of prior EM and accounting discretion is limited, the ability to use discretionary accruals to bias earnings upward decreases with the extent to which NOA are overstated by prior EM. In the context of preceding discussion, we note that overstated NOA constrains current-period EM, when it results from past accrual-based EM *and* when at least some portion of past discretionary accruals reverses back to the income statement. Although both conditions must hold, the second condition is more important, as an overstated NOA balance, regardless of magnitude, does not restrict earnings management unless it reverses.

To illustrate the second condition, assume that in period t, two identical firms accelerate sales, except that firm A recognizes $100 million sales three quarters early, and firm B recognizes $30 million one quarter early. That is, firm A's discretionary accruals reverse in period t+3, and firm B's discretionary accruals reverse in period t+1. Thus, although firm A's period t+1 beginning NOA exceeds that of firm B, firm B, but not firm A, faces an EM constraint in period t+1. To illustrate the first condition (that an overstated NOA balance results from past accrual-based EM), consider a "real" earnings management scenario, in which the firm boosts sales revenue by offering lenient credit terms. Such behavior, which inflates the firm's receivable balance in relation to sales for the period (and thus, inflates NOA), does not constrain future EM, as the inflated receivables are collected (a balance-sheet-to-balance-sheet entry), rather than reversing to the income statement (a balance-sheet-to-income-statement entry). That is, the inflated receivables impose no constraint on future EM because a reversal of discretionary accruals back to the income statement does not occur in this case.

A second constraint can arise when past cumulative reporting bias (CDA_{t-1}) reduces the ability to introduce new d_t^{new} in the same direction. More specifically, because accounting discretion is limited, past "consumption" of income-increasing (income-decreasing) accounting methods reduces the portfolio of firm's income-increasing (income-decreasing) accounting choices in the future (Barton and Simko 2002; Fan et al. 2010).[5] Further, a large NOA balance can alert auditors or security analysts to possible EM. Such effects can manifest as an inverse association between the NOA balance and the propensity to meet or beat earnings benchmarks.

A Model of the Discretionary Accrual Reversal Process

Reversals of discretionary accruals can take many forms, making formal analysis that captures all aspects of discretionary accrual reversal complex and intractable. Hence, to model the consequences of reversals, we impose two simplifying assumptions about the properties of discretionary accruals. We designate discretionary accruals initiated in period t as d_t. The first assumption is:

> **Assumption 1:** Discretionary accrual initiated in period t (d_t) reverses in its entirety in period $t+n$, where $n \geq 1$.

Denote $1/n$ ($n = 1, 2, 3, \ldots$) as the "speed" of discretionary accrual reversal such that greater $1/n$ (lower n) implies faster reversal.[6] For example, accelerating revenue recognition by one quarter implies $n = 1$, and we say that reversal speed is 1. On the other hand, discretionary accruals for items typically settled later—for example, reserves for warranty commitments—have $n > 1$, which implies a slower reversal speed—that is, a reversal speed less than 1. Following Expression (1), Assumption 1 implies:

---

[5] Fan et al. (2010) hypothesize and find that managers use classification shifting to inflate core earnings when accruals manipulation is constrained in the sense developed by Barton and Simko (2002).

[6] An alternative characterization where discretionary accruals reverse uniformly over the subsequent n periods is more complex algebraically, but yields empirical predictions consistent with those developed in this section without offering incremental insights.

$$X_t = X_t^* + (d_t - d_{t-n}) \tag{1a}$$

where the EM constraint for period t is d_{t-n}.

To consider the balance sheet effects of earnings management, recall that double-entry bookkeeping under accrual accounting implies that NOA are inflated when earnings are overstated. In particular, Assumption 1 implies that cumulative bias in NOA at the beginning of a period, CDA_{t-1}, is the sum of discretionary accruals initiated during the prior n periods: $CDA_{t-1} = d_{t-1} + d_{t-2} + d_{t-3} + \ldots + d_{t-n} = \sum_{i=1}^{n} d_{t-i}$. The second assumption facilitates tractable analysis of CDA.

Assumption 2: Discretionary accruals (d_t) are identically distributed as (μ, σ^2).

Assumption 2 means that d_t can be expressed as $d_t = \mu + \varepsilon_t$, where ε_t has zero mean and variance σ^2. If $d_t > 0$, then a firm introduces an optimistic bias to overstate period t net income; if $d_t < 0$, then the firm introduces a pessimistic bias to understate period t net income. Assumption 2 also accommodates an over-time bias in financial reporting, where $\mu > 0$ ($\mu < 0$) implies that the firm consistently introduces an optimistic (pessimistic) bias in prior earnings. If reporting is consistently optimistic over an extended period ($\mu > 0$), then the cumulative bias in NOA (i.e., CDA) is greater for firms with slower reversal speed ($1/n$). That is, $CDA_{t-1} = \sum_{i=1}^{n} d_{t-i} = n\mu + \sum_{i=1}^{n} \varepsilon_{t-i}$ varies directly with n.[7]

We do not require the ε_t's to be independently distributed. To the contrary, given the incentive to manage earnings in one direction on a consistent basis, we presume that ε_t's (and thus, d_t's) are either positively autocorrelated or uncorrelated over time. In particular, $(\zeta_k) = \theta^k$ where $0 \leq \theta < 1$ and ζ_k is the kth order autocorrelation coefficient of ε_t.

Finally, to establish a link between the income statement and the balance sheet, and to connect with the framework of Barton and Simko (2002), observe that period t EM constraint d_{t-n} is a fraction ($1/n$) of the cumulative discretionary accruals at the beginning of the period (CDA_{t-1}). In particular, Assumptions 1 and 2 imply:

$$\begin{aligned} EM\ constraint(d_{t-n}) &= \left(\frac{d_{t-n}}{\sum_{i=1}^{n} d_{t-i}}\right)\sum_{i=1}^{n} d_{t-i} = \left(\frac{d_{t-n}}{\sum_{i=1}^{n} d_{t-i}}\right) CDA_{t-1} = \left(\frac{\mu + \varepsilon_{t-n}}{n\mu + \sum_{i=1}^{n} \varepsilon_{t-i}}\right) CDA_{t-1} \\ &= \frac{1}{n}\left[\frac{\left(n\mu + \sum_{i=1}^{n} \varepsilon_{t-i}\right) + \left(n\varepsilon_{t-n} - \sum_{i=1}^{n} \varepsilon_{t-i}\right)}{n\mu + \sum_{i=1}^{n} \varepsilon_{t-i}}\right] CDA_{t-1} \\ &= \left(\frac{1}{n}\right) CDA_{t-1} + \left(\varepsilon_{t-n} - \frac{1}{n}\sum_{i=1}^{n} \varepsilon_{t-i}\right) \\ &= \left(\frac{1}{n}\right) CDA_{t-1} + \psi_t \end{aligned} \tag{2}$$

where $\psi_t = \varepsilon_{t-n} - \frac{1}{n}\sum_{i=1}^{n} \varepsilon_{t-i}$ with mean equal to 0.

---

[7] In the case of income smoothing that takes the form of alternately income-increasing (e.g., $\mu_{t-1} < 0$) and income-decreasing discretionary accruals (e.g., $\mu_t > 0$), NOA on average is unbiased, and the prediction holds trivially.

Expression (2) demonstrates that the EM constraint increases with reversal speed ($1/n$), *conditional* on the beginning balance of cumulative discretionary accruals (CDA_{t-1}). Our hypothesis, stated in the null form, is as follows:

Hypothesis: Conditional on the magnitude of cumulative bias in NOA (CDA), the earnings management constraint is unrelated with the speed of discretionary accrual reversal.

The alternative hypothesis is that, given prior earnings management that yields non-zero CDA, the ability to manage earnings toward income targets *decreases* with the reversal speed of discretionary accruals initiated in prior periods. Rejecting the null in favor of the alternative hypothesis implies that net operating assets and discretionary accrual reversal speed need to be considered *jointly* when specifying the probability of meeting or beating earnings expectations, designated *Prob (MBE)*. Thus, to evaluate the incremental contribution of reversal speed as a determinant of EM, we address the following empirical specification that extends Barton and Simko (2002).[8]

$$Prob\ (MBE) = \gamma_0 + \gamma_1 CDA + \gamma_2(1/n * CDA) + \gamma_3(1/n) + \ldots + e_{i,t}, \tag{3}$$

where $1/n$ is the reversal speed of discretionary accruals.

Consider how this characterization reconciles with the Barton and Simko (2002) characterization. Barton and Simko (2002) link *NOA* (the empirical proxy for CDA) with the probability of achieving earnings objectives, but do not explicitly consider reversal speed. They interpret $\gamma_1 < 0$ as evidence that the ability to optimistically bias current earnings decreases with the magnitude of the optimistic bias embedded in CDA. Our analysis does not contradict the relevance of NOA, but it indicates that reversal speed needs to be considered with NOA jointly to assess the consequences of discretionary accrual reversal. More specifically, the bias in NOA at the beginning of the current period that results from prior earnings management (CDA) affects the ability to manage period t reported earnings directly when prior discretionary accruals reverse. Thus, we focus on the interaction between CDA and the reversal speed, $1/n * CDA$. In particular, a negative estimate γ_2 on the interaction supports a characterization where the constraint on period t EM caused by prior EM varies directly with the speed of discretionary accrual reversal.

At least two additional points are relevant for designing empirical tests of the hypothesis. First, because reversal speeds are unobservable, the investigation requires empirical estimates. When constructing such estimates, we expect systematic differences across firms, which depend on, among other factors, the costs and benefits of using specific accounts to manage earnings and cross-sectional differences in business cycles and working capital management policies. This point suggests the use of firm-specific estimates of reversal speeds.

Second, we anticipate within-firm differences in the accounts used to initiate discretionary accruals. In particular, current balance sheet items—e.g., accounts receivable, accounts payable, short-term accrued expenses, and inventories—are, according to GAAP, expected to be settled within one fiscal year. Thus, if managers stay within the boundary of GAAP, then discretionary accruals created using current accounts typically reverse within four quarters.[9] In contrast, non-current discretionary accruals—e.g., depreciation, amortization of deferred expenditures, warranty liabilities, and reserves—reverse more slowly. Hence, we distinguish current from non-current balance sheet accounts, both when estimating the reversal speed and when constructing proxies for cumulative biases from prior earnings management.

---

[8] Barton and Simko (2002) examine ten one-cent intervals above and below the earnings target. We do not use this approach because the reversal of discretionary accruals in period t occurs regardless of the period t income objective.

[9] Dechow and Dichev (2002) report that all firms in their sample have operating cycles of less than one year.

III. RESEARCH DESIGN

Measures of Discretionary Accrual Reversal Speed

Recall that most empirical measures of discretionary accruals are "net" discretionary accruals (i.e., $NDA_t = d_t - d_{t-n}$), not new discretionary accruals initiated in the current period (d_t).[10] Procedures to isolate the reversed component from net discretionary accruals are not well developed in the literature. As a result, we rely on the widely used Jones (1991) model and utilize the autocorrelation structure of estimated residuals from the Jones model (i.e., *NDA*) to construct two proxies for the reversal speed.

More specifically, denoting ρ_k as the *k*th order autocorrelation coefficient of NDA_t, the following proposition underlies the first measure (see Appendix A for proof).

> **Proposition 1:** If period *t* discretionary accruals (d_t) reverse fully in period $t+n$, where $n \geq 1$, then the minimum value of *k*th order autocorrelation ρ_k is achieved when $k = n$.

Simply stated, if discretionary accruals (d_t) reverse only and fully in the *n*th period (Assumption 1), then the *n*th-order autocorrelation ρ_n for *NDA* (the Jones model residuals) is the most negative among all orders of autocorrelation coefficients. Therefore, the most negative autocorrelation coefficient is used to infer reversal speed $1/n$. For example, given a time-series of quarterly net discretionary accruals for a particular firm, if the third-order autocorrelation coefficient is the smallest (most negative) among all orders of autocorrelation, then $n = 3$ (i.e., speed $= 1/3$), which means that discretionary accruals initiated in quarter *t* reverse in quarter $t+3$.

The second measure is the negative of the first-order correlation of *NDA*. Reasoning that supports this measure is summarized by the following proposition (see Appendix A for proof).

> **Proposition 2:** The first-order autocorrelation coefficient of NDA_t (ρ_1) varies inversely with the reversal speed of discretionary accruals (d_t).

The intuition behind this proposition is that the first-order autocorrelation coefficient is most negative when the reversal occurs in the next period ($n = 1$). As accruals reverse more slowly ($n > 1$), the first-order autocorrelation coefficient becomes less negative. We reverse the sign of the first-order autocorrelation (which is typically negative) to obtain a measure ($-\rho_1$) that varies directly with reversal speed.

The first reversal speed measure ($1/n$) is discrete, whereas the second ($-\rho_1$) is a continuous measure. The second measure relies solely on the first-order autocorrelation coefficient of *NDA*—and, thus, requires a shorter time-series to estimate—whereas the first measure considers the autocorrelation structure of *NDA* more completely. Moreover, the first measure $1/n$ is predicated on, and therefore follows more directly from, the analysis that yields Expression (2). Details of the process that yields these measures are presented below.

Observe that the proposed measure of reversal speed captures systematic, rather than random, patterns in discretionary accrual creation and reversal. To the extent that no firm-specific, systematic pattern in discretionary accruals exists, reversal speed estimates constitute random noise. Our analysis is, therefore, a joint test for the construct validity of the cumulative discretionary accrual (CDA) measure and the reversal speed estimate.

Estimation of Discretionary Accrual Reversal Speed

The approaches described above to estimate reversal speeds require time-series of net discretionary accruals (*NDA*). To obtain *NDA* for both current and non-current accounts, we modify the Jones model to accommodate quarterly data (Rangan 1998; Baber et al. 2006). In particular, for current discretionary accruals, we estimate:

---

[10] In particular, Jones (1991) model residuals are estimates of net discretionary accruals NDA_t, as the dependent variable is total accruals (or a *change* in the NOA balance).

$$\frac{ACC\_WC_{it}}{A_{it-1}} = a_{0t} + a_{1t}\frac{1}{A_{it-1}} + a_{2t}\frac{\Delta Sales_{it}}{A_{it-1}} + \tau_{it}. \tag{4}$$

Variables are computed as in the extant literature (e.g., Richardson et al. 2005). Specifically, current accruals ACC_WC_{it} is the t–1 to t change[11] in non-cash working-capital accounts for firm i (Compustat quarterly data items $\Delta ACTQ - \Delta CHEQ - \Delta LCTQ + \Delta DLCQ$).[12] $\Delta Sales_{it}$ is the quarter t–1 to t change in sales revenue, and A_{it-1} is total assets at the end of quarter t–1. We estimate Expression (4) cross-sectionally for each industry-year-quarter combination, where industries are delineated using two-digit SIC codes.[13] Regression residuals τ_{it} from Expression (4) are designated net current discretionary accruals for firm i and quarter t.

We follow a similar procedure for non-current accruals. In particular, we estimate the following cross-sectional specification for each industry-year-quarter combination.

$$\frac{ACC\_NC_{it}}{A_{it-1}} = b_{0t} + b_{1t}\frac{1}{A_{it-1}} + b_{2t}\frac{\Delta Sales_{it}}{A_{it-1}} + b_{3t}\frac{PPE_{it}}{A_{it-1}} + v_{it}. \tag{5}$$

The dependent variable, non-current accrual (ACC_NC_{it}), is the difference between working capital accruals ACC_WC_{it}, and total accruals (ACC_{it}) computed as the change in net operating assets ($\Delta SEQQ - \Delta CHEQ + \Delta DLCQ + \Delta DLTTQ$).[14] That is, $ACC\_NC_{it} = ACC_{it} - ACC\_WC_{it}$. PPE_{it} is gross property, plant, and equipment at the end of quarter t. Residuals v_{it} from Expression (5) are net non-current discretionary accruals.

To identify the autocorrelation structure for each time-series, we estimate

$$NDA_{it} = \alpha_{it} + \rho_k NDA_{it-k} + c_{1,i}Q1 + c_{2,i}Q2 + c_{3,i}Q3 + e_{it}. \tag{6}$$

The slope coefficient ρ_k is the kth order autocorrelation coefficient, and $Q1$, $Q2$, $Q3$ are indicator variables that distinguish each of the first three fiscal quarters.

As discussed previously, the first measure of reversal speed ($1/n$) is computed using the most negative autocorrelation (ρ_k), based on firm-specific estimation of Expression (6). GAAP accounting prescribes that accruals classified as current (as opposed to non-current) convert to cash, or be otherwise extinguished, within the normal business operating cycle. Hence, for current discretionary accruals, we specify n^{ST} as the order that yields the lowest ρ_k among the first four orders.[15] For non-current discretionary accruals, we specify n^{LT} as the order that yields the lowest ρ_k of the first 16 orders.[16] The second measure ($-\rho_1$) uses the firm-specific estimate of the first-order autocorrelation from Specification (6).

---

[11] The first-difference, not seasonal difference, of NOA must be used to estimate quarter t discretionary accruals. Seasonal differencing of NOA yields an estimate of net discretionary accruals for the preceding four quarters combined.

[12] $ACTQ$ is total current assets; $CHEQ$ is cash and short-term investments; $LCTQ$ is total current liabilities; and $DLCQ$ is debt in current liabilities. These four variables are Compustat data items 40, 36, 49, and 45, respectively.

[13] Using firm-by-firm estimates of ACC_WC_{it} yield comparable empirical results.

[14] $SEQQ$ is shareholders' equity, and $DLTTQ$ is long-term debt. These variables are Compustat data items 60 and 51, respectively.

[15] Results are comparable when we choose the lowest among the first eight orders of autocorrelation coefficients of current NDA, or among the first 16 orders. As expected, the significance levels are progressively weaker as maximum n increases.

[16] Reversal cycles of non-current accruals are likely longer than four years (16 quarters). From a practical standpoint, however, increasingly fewer observations are available to estimate autocorrelations as the estimation period increases. Thus, we restrict the estimated reversal speed for long-term discretionary accruals, as a trade-off between accuracy and sample size.

Estimation of CDA (Cumulative Discretionary Accruals)

The empirical analysis also requires measures of the cumulative effect of prior earnings management (CDA). Following Barton and Simko (2002), we presume that CDA is indicated by net operating assets at the beginning of the period (NOA_{t-1}).

Consistent with procedures employed to estimate reversal speed, we partition NOA into current and non-current NOA.[17] We calculate current net operating assets, designated *NOA_WC*, as Compustat quarterly data items ($ACTQ - CHEQ - LCTQ + DLCQ$) deflated by sales.[18] We calculate non-current net operating assets, designated *NOA_NC*, as the difference between total net operating assets (Compustat quarterly data items $SEQQ - CHEQ + DLCQ + DLTTQ$) and current *NOA_WC* computed above.

DeFond (2002) notes that raw measures of net operating assets are potentially compromised by industry- and firm-specific differences unrelated to net asset overstatements. If so, then controlling for "normal" levels of net operating assets can reduce noise and increase statistical power. Thus, we employ time-, size-, and industry-adjusted NOA as a measure of CDA. In particular, we partition firms in each two-digit industry into ten size-deciles for each year, and then estimate the mean *NOA_WC* and mean *NOA_NC* for each decile. Empirical proxies for current (non-current) CDA are *NOA_WC (NOA_NC)* net of the corresponding mean effects of industry, firm size, and firm year.[19]

Specifications of the Probability of Meet or Beat Earnings Expectations

Incentives to meet or beat analyst forecasts are particularly strong (Burgstahler and Dichev 1997; Bartov et al. 2002; Matsumoto 2002), so we test the hypothesis by investigating the ability to meet forecasts or report positive earnings surprises. In particular, we classify observations according to whether reported earnings meet or beat analyst forecasts. The dependent (indicator) variable is set to unity when I/B/E/S actual EPS equals or exceeds the most recent consensus quarter t forecast prior to the earnings announcement ($MBE = 1$), and 0 otherwise.[20]

Recall that we consider the reversals of current and non-current discretionary accruals separately. The complete empirical specification is as follows:

$$\begin{aligned}\ln\left[\frac{\Pr(MBE_t = 1)}{\Pr(MBE_t = 0)}\right] &= \beta_0 + \beta_1(NOA\_WC_{it-1}) + \beta_2(NOA\_WC_{it-1} * speed^{ST}_{it-1}) \\ &+ \beta_3(NOA\_NC_{it-1}) + \beta_4(NOA\_NC_{it-1} * speed^{LT}_{it-1}) \\ &+ \beta_5(speed^{ST}_{it-1}) + \beta_6(speed^{LT}_{it-1}) + Controls + \varepsilon_{it},\end{aligned} \tag{7}$$

where *NOA_WC* and *NOA_NC*, respectively, designate industry- and size-adjusted current and non-current NOA at the beginning of the quarter, and $speed_{it-1}$ is either of the discretionary accrual reversal speed measures described above. Superscripts *ST* and *LT* distinguish the reversal speeds for current (*ST*) versus non-current (*LT*) discretionary accruals.

Negative parameter estimates on the interaction terms—i.e., negative β_2 and β_4—are consistent with the hypothesis that the earnings management constraint increases with the reversal speed of current (or non-current) discretionary accruals, *holding constant* the bias in current (or non-current)

---

[17] Barton and Simko (2002) partition NOA into three components, i.e., working capital, fixed assets, and other long-term assets.

[18] Choice of the deflator follows Barton and Simko (2002). Results are comparable when the book value of total assets is the deflator.

[19] Barton and Simko (2002) use NOA without adjusting for year, industry, and size. Results (untabulated) are similar, but slightly weaker, for raw (unadjusted) NOA measures.

[20] Results are comparable when we use the median forecast issued within 45 days of the earnings announcement.

NOA. Expectations regarding parameter estimates on the main effects *NOA_WC* and *NOA_NC* are not straightforward when the interaction terms are included. If mechanistic reversal of prior discretionary accruals is the only constraint on the ability to manage current earnings, then we expect parameters on these main effects to be zero. On the other hand, if prior earning management undermines the ability to manage earnings indirectly—for example, if a large NOA itself directs attention of auditors or security analysts to the possibility of earnings management—then we expect a negative parameter on NOA. Because we are uncertain about whether NOA matters only directly—that is, only through mechanistic reversal of discretionary accruals—we predict parameters on the main effects NOA to be zero or negative (i.e., $\beta_1 \leq 0$ and $\beta_3 \leq 0$). Including reversal speed as a main effect ensures that results for the interaction cannot be attributed to unspecified main effects; we offer no predictions about the signs of β_5 and β_6.

Finally, control variables (*Controls*) are identical to those used in Barton and Simko (2002). Specifically, for quarter *t*:

$SHARES_t$ = weighted average common shares outstanding;
$BIG5_t$ = 1 if the firm has a Big 5 auditor, 0 otherwise;
MTB_t = market value of common $shares_t$/shareholders' $equity_t$;
$LTGN\_RISK_t$ = 1 if the industry is pharmaceuticals/biotechnology (SIC codes 2833–2836, 8731–8734), computers (3570–3577, 7370–7374), electronics (3600–3674), or retail (5200–5961), 0 otherwise;
$ANALYSTS_t$ = number of analysts issuing an EPS forecast;
$PREV\_MB_t$ = 1 if reported earnings meet or exceed the consensus forecast in the prior quarter, 0 otherwise;
$CV\_FORECAST_t$ = coefficient of variation (i.e., standard deviation/mean) in analysts' most recent forecasts for quarter *t*;
$DOWN\_REV_t$ = 1 if at least one of the firm's analysts revised his or her forecast down prior to the end of quarter *t* but after the earnings announcement date for quarter *t*–1, 0 otherwise;
$SALES\_GRW_t = (sales_t/sales_{t-4}) - 1$;
ROE_t = net $income_t$/shareholders' $equity_t$;
$\Delta ROE_t = ROE_t - ROE_{t-1}$; and
MKT_CAP_t = *ln*(market value of common $shares_t$).

IV. SAMPLE

Sample Selection

Quarterly earnings surprise measures are computed as reported EPS less the most recent consensus analyst forecasts from the 1993 to 2007 I/B/E/S unadjusted actual and summary files. I/B/E/S changes the reporting of actual earnings per share in 1993 (Abarbanell and Lehavy 2002), such that beginning the sample period in 1993 facilitates consistent computation of earnings surprises. We delete firm-quarter observations with missing control variables, and exclude utilities and financial services firms (two-digit SIC codes 49 and 60–67). These criteria yield 129,323 firm-quarter observations for 7,424 firms.

For each firm-year-quarter observation, we estimate reversal speed using a rolling window of net discretionary accruals NDA, obtained from Specification (4) or (5), for the 12 years (48 quarters) ending the prior quarter. For each 12-year estimation period, we impose two requirements: first, we require at least 20 quarterly NDA observations to estimate each of the first 16 orders of autocorrelation coefficients. Second, we ensure that the firm has no significant mergers or acquisitions (M&A), defined as annual acquisitions (Compustat annual data item *AQC*) greater than 20 percent of the firm's beginning-of-the-year assets. The latter requirement is necessary because

M&A activities compromise the structure of the reversal process and the stationarity of time-series estimates.[21] The first requirement reduces the sample to 59,841 firm-quarter observations (3,100 firms); the second requirement reduces the final sample to 36,210 firm-quarter observations (2,109 firms). Thus, data requirements to estimate reversal speeds yield a sample smaller than the 3,649 firms used in Barton and Simko (2002).

Estimation of Reversal Speed

Table 1 reports the sample means and medians for the first 16 orders of autocorrelations computed for current and non-current NDA. As shown in columns 1 and 2, mean (median) autocorrelation of the first order, −0.160 (−0.154), is the most negative among the 16 autocorrelations for current NDA. For non-current NDA (columns 5 and 6), the most negative mean and median autocorrelation occur after the eighth order. Moreover, juxtaposing columns 1 and 2 with columns 5 and 6 indicates that mean and median autocorrelations are typically more negative for current than for non-current discretionary accruals. Thus, consistent with intuition, the autocorrelation structure for current discretionary accruals differs from the structure for non-current discretionary accruals.

For *current* discretionary accruals, we present the relative frequencies that the *k*th-order autocorrelation is the most negative among the first 16 orders in column 3. We show the relative frequencies among the first four lags in column 4. Column 7 shows the corresponding frequency for non-current discretionary accruals among the first 16 orders. Comparing column 3 and column 7 indicates that, for 41 percent of the firm-quarter observations, the most negative autocorrelation of current discretionary accruals occurs within the first four orders (22 percent + 10 percent + 7 percent + 2 percent), which is a much higher frequency than the 17 percent (5 percent + 4 percent + 4 percent + 4 percent) for non-current discretionary accruals. These comparisons confirm the expectation that current discretionary accruals reverse more quickly than non-current discretionary accruals.[22]

Finally, recall that we use two reversal speed measures. The first is $1/n$, where n corresponds to the most negative of the first four autocorrelation coefficients displayed in column 4. Thus, $n = 1$ for 43 percent of the observations; $n = 2$ ($1/n = 0.5$) for 29 percent of the observations; $n = 3$ ($1/n = 0.333$) for 21 percent of the observations; and $n = 4$ ($1/n = 0.25$) for 7 percent of the observations. We obtain a substantially different result for non-current accruals (column 7), however. In particular, $n = 15$ or 16 (9 percent each) is the most frequent order with the lowest autocorrelation of the first 16 autocorrelations for non-current accruals. These estimates confirm that non-current discretionary accruals reverse more slowly than current discretionary accruals. The mean and median of the second reversal speed measure are in the first row of Table 1.

Table 2, Panel A reports summary statistics of variables used to estimate Expression (7). The last column shows Spearman correlations with the measure of earnings surprise, *MBE*. The focus of

---

[21] The 20 percent cutoff is determined by examining the distribution of acquisition/assets of all Compustat firm-year observations, and balancing data preservation against significance of M&A activities. Ninety-six percent of all Compustat firm-years report acquisitions less than 20 percent, and therefore, the sample includes firms with small M&A amounts less than the 20 percent cutoff. Comparable results obtain when 15 percent or 25 percent cutoffs are used.

[22] This point, coupled with comparisons of columns (3) and (4), raises questions about why estimates of reversal periods for current discretionary accruals exceed four. We conjecture that the explanation has to do with seasonality. More specifically, current NDA four quarters apart are significantly positively related—i.e., mean fourth-order autocorrelation is 0.194 (two-tailed $p < 0.01$). Moreover, this fourth-order autocorrelation coefficient is greater on average than all autocorrelation coefficients beyond the fourth order. Investigating this point further is beyond the scope of the study, but we note that results are similar qualitatively when we use the first eight or 16 orders of autocorrelations to compute current reversal speeds.

TABLE 1

Distributions of Autocorrelation Coefficients of Net Discretionary Accruals (*NDA*)

	Current Net Discretionary Accruals				Non-Current Net Discretionary Accruals		
Order	**Mean (1)**	**Median (2)**	**% Most Negative of the First 16 Orders (3)**	**% Most Negative of the First 4 Orders (4)**	**Mean (5)**	**Median (6)**	**% Most Negative of the First 16 Orders (7)**
1	−0.160	−0.154	22%	43%	0.072	0.054	5%
2	−0.042	−0.034	10%	29%	0.068	0.054	4%
3	−0.092	−0.074	7%	21%	0.043	0.033	4%
4	0.194	0.163	2%	7%	0.074	0.058	4%
5	−0.093	−0.078	6%		0.003	−0.003	5%
6	−0.047	−0.038	8%		−0.006	−0.008	6%
7	−0.089	−0.072	6%		−0.017	−0.014	6%
8	0.172	0.137	2%		0.017	0.011	5%
9	−0.083	−0.068	5%		−0.035	−0.031	6%
10	−0.042	−0.034	7%		−0.032	−0.028	8%
11	−0.078	−0.067	5%		−0.036	−0.034	7%
12	0.153	0.120	2%		−0.001	−0.009	6%
13	−0.075	−0.061	5%		−0.040	−0.036	8%
14	−0.045	−0.039	6%		−0.037	−0.033	8%
15	−0.066	−0.052	5%		−0.035	−0.029	9%
16	0.141	0.110	2%		−0.005	−0.005	9%
Sum			100%	100%			100%

Current and non-current net discretionary accruals are residuals from estimating Expressions (4) and (5), respectively.

our study is the relation between *MBE* and the interaction *NOA* * *speed*. Consistent with the alternative hypothesis, *MBE* is negatively correlated with interactions of current discretionary accrual reversal speed measures and industry-size-adjusted current net operating assets, $(1/n^{ST} * NOA\_WC)$ and $(-\rho_1^{ST} * NOA\_WC)$. For the corresponding non-current accruals, one interaction $(1/n^{LT} * NOA\_NC)$ varies inversely with earnings surprise, but the other $(-\rho_1^{LT} * NOA\_NC)$ is positively related with earnings surprise.

The analysis in Section II implies a negative correlation between reversal speed and the size of the CDA balance—that is, *ceteris paribus*, the faster the speed of discretionary accrual reversal, the smaller is the cumulative bias on the balance sheet (CDA). Panel B of Table 2 shows correlations between reversal speed estimates, $(1/n)$ and $(-\rho_1)$, and CDA measures (raw and adjusted NOA). Observe that both reversal speed measures are significantly and negatively correlated with either CDA measures for current accruals, which validates the conceptual analysis in Section II. The inverse correlation does not always hold for non-current accruals, however. Finally, untabulated results show that Pearson (Spearman) correlation coefficients between the two alternative measures of current reversal speed are 0.638 (0.633) and between two measures of non-current reversal speed are 0.408 (0.253). Correlations between measures of current and non-current reversal speed are less than 0.10.

TABLE 2
Summary Statistics

Panel A: Summary Statistics

Independent Variable	**Mean**	**Std.**	**Min**	**Q1**	**Median**	**Q3**	**Max**	**Pred. Sign**	**Spearman Rank Corr. w/MBE_t**
Primary Variables									
NOA_WC_{t-1}	0.072	0.466	−1.159	−0.195	0.036	0.317	1.201	−	−0.071***
NOA_NC_{t-1}	−0.510	1.605	−4.127	−1.299	−0.522	0.070	5.454	−	−0.036***
$\left(1/n^{ST}_{t-1}\right)$	0.664	0.303	0.250	0.333	0.500	1	1	?	0.007
$\left(1/n^{LT}_{t-1}\right)$	0.175	0.211	0.063	0.071	0.100	0.167	1	?	−0.005
$NOA\_WC_{t-1} * \left(1/n^{ST}_{t-1}\right)$	0.041	0.336	−1.037	−0.112	0.019	0.181	1.096	−	−0.068***
$NOA\_NC_{t-1} * \left(1/n^{LT}_{t-1}\right)$	−0.096	0.328	−1.686	−0.162	−0.055	0.007	1.057	−	−0.031***
$\left(-\rho^{ST}_{1,t-1}\right)$	0.160	0.200	−0.342	0.028	0.154	0.295	0.643	?	−0.001
$\left(-\rho^{LT}_{1,t-1}\right)$	−0.072	0.201	−0.623	−0.189	−0.054	0.059	0.379	?	−0.003
$NOA\_WC_{t-1} * \left(-\rho^{ST}_{1,t-1}\right)$	0.001	0.115	−0.415	−0.035	−0.000	0.036	0.394	−	−0.043***
$NOA\_NC_{t-1} * \left(-\rho^{LT}_{1,t-1}\right)$	0.049	0.339	−1.093	−0.059	0.006	0.119	1.448	−	0.012**
Control Variables									
$SHARES_t$	137.104	292.419	1.168	16.241	37.285	104.400	2,107.250	−	0.139***
MTB_t	3.247	5.384	−52.774	1.526	2.360	3.798	76.087	+	0.161***
$BIG5_t$	0.923	0.266	0	1	1	1	1	?	0.057***
$LTGN\_RISK_t$	0.275	0.447	0	0	0	1	1	−	0.036***
$ANALYSTS_t$	7.541	6.708	1.000	2.000	5.000	11.000	44.000	+	0.164***
$PREV\_MB_t$	0.673	0.469	0	0	1	1	1	+	0.246***
$CV\_FORECAST_t$	0.046	0.616	−29.000	0.000	0.031	0.082	30.000	−	−0.001
$DOWN\_REV_t$	0.322	0.467	0	0	0	1	1	?	−0.081***
$SALES\_GRW_t$	0.152	0.576	−1.000	0.004	0.082	0.195	14.670	+	0.190***
ROE_t	0.021	0.199	−3.083	0.009	0.039	0.051	2.721	+	0.249***
ΔROE_t	−0.002	0.319	−4.471	−0.015	0.0001	0.012	5.520	+	0.197***
MKT_CAP_t	556.232	4,346.300	0.103	5.488	6.819	8.602	61,276.350	+	0.150***

Panel B: Correlations between *NOA* and Reversal Speed Measures

	1/*n*		$-\rho_1$	
	Pearson	**Spearman**	**Pearson**	**Spearman**
Correlations with the reversal speed of current discretionary accruals				
NOA_WC (raw)	−0.136***	−0.148***	−0.201***	−0.210***
NOA_WC (industry-sized adjusted)	−0.047***	−0.047***	−0.109***	−0.098***
Correlations with the reversal speed of non-current discretionary accruals				
NOA_NC (raw)	0.021***	−0.019***	−0.004	−0.037***
NOA_NC (industry-sized adjusted)	0.002	−0.016***	0.036***	0.015***

** and *** Denote two-tailed significance at the 0.05, and 0.01 levels, respectively.

The sample is 36,210 firm-quarters for 2,109 firms with complete 1993–2007 data in both Compustat and I/B/E/S. Utilities and financial services firms (two-digit SIC codes 49 and 60–67) are excluded.

(continued on next page)

TABLE 2 (continued)

Variable Definitions:

Dependent Variables:

$MBE = 1$, if I/B/E/S actual EPS equals or exceeds the consensus forecast for quarter t, and 0 otherwise;

Primary Independent Variables (Compustat quarterly data item numbers in brackets):

$WC_{t-1} = (ACTQ[40] - CHEQ[36] - LCTQ[49] + DLCQ[45])_{t-1}/(SALEQ[2])_{t-1}$, i.e., lagged non-cash working capital scaled by lagged sales;

$NOA_{t-1} = (SEQQ[60] - CHEQ[36] + DLCQ[45] + DLTTQ[51])_{t-1}/(SALEQ[2])_{t-1}$, lagged total net operating assets scaled by lagged sales;

$NC_{t-1} = NOA_{t-1} - WC_{t-1}$;

$NOA\_WC_{t-1} = WC_{t-1}$ – mean WC_{t-1} for the firm's partition by year, two-digit industry, and size-decile;

$NOA\_NC_{t-1} = NC_{t-1}$ – mean NC_{t-1} for the firm's partition by year, two-digit industry, and size-decile;

$(1/n_{ST})$ = 1/the number of quarters it takes for current discretionary accruals to reverse, where $n_{ST} \leq 4$;

$(1/n_{LT})$ = 1/the number of quarters it takes for non-current discretionary accruals to reverse $n_{LT} \leq 16$;

$(-\rho^{ST}_{1,\,t-1})$ = first-order autocorrelation for current net discretionary accruals;

$(-\rho^{LT}_{1,\,t-1})$ = first-order autocorrelation for non-current net discretionary accruals;

Control Variables:

$SHARES_t$ = weighted average number of common shares outstanding in quarter t;

$BIG5_t = 1$ if the firm has a Big 5 auditor in quarter t, 0 otherwise;

MTB_t = market value of common shares_t/shareholders' equity_t, both at the end of quarter t;

$LTGN\_RISK_t = 1$ if the firm is in one of the following industries: pharmaceuticals/biotechnology (SIC codes 2833–2836, 8731–8734), computers (3570–3577, 7370–7374), electronics (3600–3674), or retail (5200–5961), 0 otherwise;

$ANALYSTS_t$ = number analysts in the I/B/E/S consensus EPS forecast for quarter t;

$PREV\_MB_t = 1$ if the firm reported a nonnegative earnings surprise in quarter $t-1$ based on I/B/E/S, 0 otherwise;

$CV\_FORECAST_t$ = coefficient of variation (i.e., standard deviation/mean) in analysts' most recent forecasts for quarter t;

$DOWN\_REV_t = 1$ if at least one of the firm's analysts revised his or her forecast down prior to the end of quarter t but after the earnings announcement date for quarter $t-1$, 0 otherwise;

$SALES\_GRW_t = (sales_t/sales_{t-4}) - 1$;

ROE_t = net income_t/shareholders' equity_t;

$\Delta ROE_t = ROE_t - ROE_{t-1}$; and

$MKT\_CAP_t = \ln(\text{market value of common shares}_t)$.

V. RESULTS

Replication of Barton and Simko (2002)

To begin, we replicate the Barton and Simko (2002) finding. In contrast with Barton and Simko (2002), we partition *NOA* into two, rather than three, components—current and non-current—and we use industry-size-adjusted *NOA-to-sales ratio*, rather than the raw *NOA-to-sales ratio* in their study, to measure the bias in NOA that results from prior accrual-based earnings management. Finally, we used standard errors clustering on firm (An 2008) to control for cross-sectional dependence for each firm observation.

Results are in Table 3, Panel A. Parameter estimates on *NOA_WC* and *NOA_NC* are –0.213 and –0.036, respectively (two-tailed $p < 0.01$), confirming the Barton and Simko (2002) result. The 1993–2007 sample goes beyond the 1993–1999 period studied by Barton and Simko (2002). Events around the passage of the Sarbanes-Oxley Act (SOX) potentially influence management's incentive and ability to use the accrual process to manage earnings. For example, evidence in Cohen et al. (2008) suggests that accrual-based earnings management increases steadily from 1987 to 2002, but then declines after 2002.

Table 3, Panel B shows estimates for *NOA_WC* and *NOA_NC* by fiscal year. Prior to 2002, the parameter estimates for current net operating assets (*NOA_WC*) are consistently negative and differ reliably from zero in all years except 1994 and 2001. Estimates are most negative in 1996 (–0.458) and 1998 (–0.447), decline gradually to –0.266 in 2003, and then become insignificant

TABLE 3

The Relation between Cumulative Bias in Net Operating Assets and Current Earnings Management

Panel A: Estimates for the Full Sample (1993–2007)

Explanatory Variables	**Parameter Estimate**	**z-statistic**
Intercept	−0.715***	−9.054
NOA_WC_{t-1}	−0.213***	−6.312
NOA_NC_{t-1}	−0.036***	−3.788
$SHARES_t$ (Coeff. × 1000)	−0.090	−1.125
MTB_t	0.019***	4.700
$BIG5_t$	0.232***	4.110
$LTGN\_RISK_t$	0.054	1.419
$ANALYSTS_t$	0.055***	14.368
$PREV\_MB_t$	0.878***	29.966
$CV\_FORECAST_t$	−0.014	−0.735
$DOWN\_REV_t$	−0.583***	−20.187
$SALES\_GRW_t$	0.225***	3.696
ROE_t	0.833***	6.775
ΔROE_t	−0.213***	−3.590
MKT_CAP_t (Coeff. × 1000)	0.009***	2.663
Likelihood ratio	3,805.441	

Panel B: Estimates for Specifications by Year

	NOA_WC_{t-1}		NOA_NC_{t-1}	
Year	**Parameter Estimate**	**z-statistic**	**Parameter Estimate**	**z-statistic**
1993	−0.280**	−1.978	0.006	0.122
1994	−0.103	−0.736	0.027	0.565
1995	−0.396***	−2.793	0.153***	2.779
1996	−0.458***	−3.377	−0.126**	−2.462
1997	−0.349***	−2.735	−0.112**	−2.497
1998	−0.447***	−3.850	−0.106***	−3.040
1999	−0.300***	−2.614	0.001	0.030
2000	−0.253**	−2.265	−0.087***	−3.014
2001	−0.169	−1.590	0.010	0.367
2002	−0.176*	−1.865	−0.046*	−1.740
2003	−0.266**	−2.454	−0.050*	−1.689
2004	−0.130	−1.409	−0.040	−1.394
2005	−0.011	−0.125	−0.019	−0.766
2006	−0.076	−0.939	−0.002	−0.104
2007	−0.053	−0.662	0.001	0.050

(continued on next page)

TABLE 3 (continued)

Panel C: Estimates for Subsamples Partitioned at 2003

Explanatory Variables	**Parameter Estimates for the 1993–2003 Period (z-statistic) (1)**	**Parameter Estimates for the 2004–2007 Period (z-statistic) (2)**	**Difference between (1) and (2) (z-statistic) (3)**
Intercept	−1.427***	−0.368***	1.058***
	(−11.967)	(−4.878)	(7.821)
NOA_WC_{t-1}	−0.278***	−0.072	0.206***
	(−6.660)	(−1.509)	(3.435)
NOA_NC_{t-1}	−0.040***	−0.020	0.020
	(−3.436)	(−1.396)	(1.123)
$SHARES_t$ (Coeff. × 100)	−0.051***	−0.001	0.050***
	(−5.730)	(−0.096)	(4.698)
MTB_t	0.014***	0.016***	0.003
	(3.004)	(3.142)	(0.399)
$BIG5_t$	0.157*	0.221***	0.064
	(1.925)	(3.218)	(0.651)
$LTGN\_RISK_t$	0.122***	0.008	−0.114
	(2.649)	(0.144)	(−1.595)
$ANALYSTS_t$	0.029***	0.055***	0.026***
	(5.434)	(9.873)	(3.675)
$PREV\_MB_t$	0.908***	0.728***	−0.181***
	(26.176)	(14.581)	(−3.047)
$CV\_FORECAST_t$	−0.020	−0.004	0.016
	(−0.840)	(−0.117)	(0.379)
$DOWN\_REV_t$	−0.604***	−0.590***	0.014
	(−17.017)	(−12.092)	(0.236)
$SALES\_GRW_t$	0.282***	0.160***	−0.122
	(2.609)	(2.800)	(−1.004)
ROE_t	0.787***	0.657***	−0.129
	(5.447)	(3.874)	(−0.599)
ΔROE_t	−0.186**	−0.179**	0.007
	(−2.360)	(−1.993)	(0.054)
MKT_CAP_t (Coeff. × 100)	15.440***	0.001**	−15.440***
	(9.301)	(2.553)	(−9.301)
n	24,285	11,925	
Likelihood ratio	2,998.031	954.848	

*, **, *** Significance at the 0.10, 0.05, and 0.01 levels (two-tailed), respectively.

Entries are parameter estimates from logit specifications of *MBE*, the probability of meeting or beating the most recent quarterly consensus analyst forecast. The sample includes 36,210 firm-quarters for 2,109 firms with complete 1993–2007 data in both Compustat and I/B/E/S. Utilities and financial services firms are excluded. z-statistics are based on standard errors clustering on firm. Other variables are defined in Table 2. In all panels, estimates for fiscal year indicator variables are not reported. In Panel B, we include control variables in these specifications, but we do not report estimates for these variables.

for all years after 2003. Although 2003 is the first year in the post-SOX period, the significantly negative coefficient of −0.266 possibly reflects current discretionary accruals initiated in 2002 reversing in 2003. Failure to achieve statistical significance after 2003 supports the Cohen et al. (2008) finding that accrual-based earnings management declines following SOX.[23] Results in Table 3, Panel C confirm that the constraint imposed on current earnings management by prior current discretionary accruals differs significantly between the pre-SOX and post-SOX periods.

Table 3, Panel B also presents annual estimates on non-current operating assets NOA_NC. Except for a significantly positive coefficient of 0.153 in 1995, most coefficients are negative, but much smaller in magnitude than corresponding coefficients on NOA_WC. Most estimates on NOA_NC are reliably negative from 1996 to 2002 (two-tailed $p < 0.10$), but estimates are insignificant after 2003. Finally, evidence in Panel C indicates that estimates for NOA_NC are not reliably less after 2003 than before. Based on the above evidence, we investigate the role of reversal speed using data for 1995 to 2003, the period where accrual-based earnings management is prevalent.[24] Results using 2004–2007 data show that neither the balance of NOA nor its interaction with the reversal speed is a significant factor in the likelihood of meeting or beating the analyst forecast.

Primary Results

Table 4 presents parameter estimates for logit specifications of Expression (7). Results based on the first set of (second) reversal speed measures, i.e., $1/n^{ST}$ and $1/n^{LT}$ ($-\rho_1^{ST}$ and $-\rho_1^{LT}$) are reported in the first (second) column. Collinearity between main effects and interactions potentially compromises statistical power (Aiken and West 1996).[25] Thus, we display in the third and fourth columns estimates for specifications where variables are centered (mean-adjusted).

The evidence for current, but not for non-current, net operating assets supports the hypothesis developed in Section II. In particular, estimates on the interaction ($NOA\_WC * speed^{ST}$) are negative and statistically significant (two-tailed $p < 0.01$) for all specifications. Thus, conditional on the magnitude of the bias in net current operating assets as measured by NOA_WC, the earnings management constraint from the reversal of prior discretionary accruals increases with the reversal speed. Thus, we reject the null hypothesis that the ability to meet or beat forecasts is unrelated to reversal speed in favor of the alternative hypothesis that reversal speed restricts the ability to achieve earnings forecast. Evidence regarding non-current discretionary accruals differs, however, as none of the coefficients on the interaction term between NOA_NC and the reversal speed of non-current discretionary accruals is statistically significant. A potential explanation for why the alternative hypothesis is supported for current, but not for non-current, NOA is that managers are better able to control the reversal of non-current discretionary accruals than the reversal of current discretionary accruals. That is, reversals of prior non-current discretionary accruals are themselves discretionary relative to reversals of prior working capital discretionary accruals.[26] A second

---

[23] We also find (untabulated) that the magnitude of the industry-size adjusted NOA_WC and NOA_NC in 2001–2002 are higher than their corresponding levels either before 2001 or after 2002. This spike in NOA is consistent with Cohen et al. (2008) who find increased accrual-based earnings management activities during the period immediately preceding SOX.

[24] The hypothesis presumes that a substantial amount of positive or negative "bank" of cumulative discretionary accruals exists (due to consistent reporting bias toward one direction). If earnings management declines significantly, then the amount of CDA is also small, and thus, little is left to reverse.

[25] Pearson correlations between NOA_WC and $NOA\_WC * speed^{ST}$ are 0.91 (0.54), and between NOA_NC and $NOA\_NC * speed^{LT}$ are 0.69 (−0.33), when speed is measured as $1/n$ ($-\rho_1$). After mean-adjusting these measures, Pearson correlations between NOA_WC and $NOA\_WC * speed^{ST}$ are −0.03 (−0.07), and between NOA_NC and $NOA\_NC * speed^{LT}$ are −0.05 (−0.04).

[26] Results are robust to procedures suggested by Norton et al. (2004) to calculate marginal effects of interaction terms in logit and probit models.

TABLE 4

Primary Results: 1993 to 2003

		Coefficient Estimates (z-statistic in parentheses)			
		Speed = 1/n	Speed = $(-\rho_1)$	Speed = 1/n	Speed = $(-\rho_1)$
Explanatory Variables	**Predicted Sign**	**Raw Measures**		**Mean-Adjusted Measures**	
Intercept	?	−1.432***	−1.421***	0.421***	0.417***
		(−11.523)	(−11.833)	(7.523)	(7.424)
NOA_WC_{t-1}	−/0	−0.064	−0.197***	−0.285***	−0.300***
		(−0.699)	(−3.972)	(−6.884)	(−7.204)
NOA_NC_{t-1}	−/0	−0.032*	−0.039***	−0.043***	−0.042***
		(−1.910)	(−3.096)	(−3.641)	(−3.607)
$NOA\_WC_{t-1} * speed^{ST}_{t-1}$	−	−0.333***	−0.675***	−0.321***	−0.645***
		(−2.657)	(−3.428)	(−2.603)	(−3.519)
$NOA\_NC_{t-1} * speed^{LT}_{t-1}$	−	−0.073	0.060	−0.055	0.075
		(−0.835)	(1.031)	(−1.042)	(1.444)
$speed^{ST}_{t-1}$	?	0.032	−0.159*	0.008	−0.202**
		(0.535)	(−1.697)	(0.142)	(−2.216)
$speed^{LT}_{t-1}$	?	−0.112	0.165*	−0.073	0.135
		(−1.130)	(1.690)	(−0.832)	(1.435)
$SHARES_t$	−	−0.001***	−0.001***	−0.001***	−0.001***
		(−5.730)	(−5.795)	(−5.730)	(−5.795)
MTB_t	+	0.014***	0.013***	0.014***	0.013***
		(2.982)	(2.904)	(2.982)	(2.904)
$BIG5_t$	?	0.152*	0.162**	0.153*	0.162**
		(1.877)	(1.972)	(1.883)	(1.977)
$LTGN\_RISK_t$	−	0.119***	0.120***	0.120***	0.119***
		(2.597)	(2.608)	(2.606)	(2.593)
$ANALYSTS_t$	+	0.029***	0.028***	0.029***	0.028***
		(5.467)	(5.483)	(5.467)	(5.483)
$PREV\_MB_t$	+	0.906***	0.903***	0.906***	0.903***
		(26.110)	(26.177)	(26.118)	(26.171)
$CV\_FORECAST_t$	−	−0.020	−0.020	−0.020	−0.020
		(−0.835)	(−0.823)	(−0.839)	(−0.823)
$DOWN\_REV_t$	?	−0.605***	−0.605***	−0.605***	−0.605***
		(−17.037)	(−17.039)	(−17.042)	(−17.088)
$SALES\_GRW_t$	+	0.284***	0.285***	0.285***	0.286***
		(2.643)	(2.631)	(2.645)	(2.642)
ROE_t	+	0.788***	0.789***	0.788***	0.789***
		(5.467)	(5.469)	(5.469)	(5.467)
ΔROE_t	+	−0.186**	−0.186**	−0.186**	−0.186**
		(−2.351)	(−2.355)	(−2.349)	(−2.351)
MKT_CAP_t	+	0.156***	0.159***	0.156***	0.159***
		(9.442)	(9.695)	(9.430)	(9.713)
Likelihood ratio		3,009.029	3,029.122	3,008.985	3,030.379

---

*, **, *** Significance at the 0.10, 0.05, and 0.01 levels (two-tailed), respectively.

Parameter estimates are from logit models of *MBE*, the probability of meeting or beating the most recent quarterly consensus analyst forecast. z-statistics are based on standard errors clustering on firm. Estimates for fiscal year indicator variables are not reported.

explanation for the insignificant results for non-current accruals is that intangibles may not be amortized, and deferred tax liabilities typically do not reverse for growing firms.

Next consider evidence regarding the main effects of net operating assets. Recall that a negative coefficient on *NOA_WC* indicates that, apart from the issue of discretionary accrual reversal, prior earnings management discourages managers to further bias earnings in the current period. An insignificant coefficient indicates that prior earnings management constrains current earnings management only through accrual reversal. Parameter estimates on *NOA_WC* are consistently negative and are statistically significant in the third and fourth columns in which all variables are centered. Thus, the results indicate that prior earnings management indirectly constrains current earnings management beyond the income effect of the mechanistic reversal of discretionary accruals. Moreover, estimates on main effect of non-current operating assets *NOA_NC* are reliably negative. Thus, estimates on the main effects of NOA continue to support the Barton and Simko (2002) characterization when discretionary accrual reversals are considered.

Further Evidence

The reversal speed measures used in Table 4 presume a linear structure among observations with different reversal speeds and, at the same time, presume a linear association between reversal speed measures and the restriction on EM attributable to discretionary accrual reversals. For example, the reversal speed measure $1/n$ ($n \leq 4$ quarters) for current discretionary accruals assumes one of four values, 1.00, 0.50, 0.33, or 0.25, such that an observation with $n = 1$ is presumed to be four times more restrictive than an observation with $n = 4$.

An alternative approach is to treat reversal speed as a discrete measure. For the first measure ($1/n^{ST}$), we use four indicator variables to classify observations according to four discrete values ($n = 1, 2, 3$, or 4). In particular, if current discretionary accruals (d_t) reverse in quarter t+n, where n = (1, 2, 3, 4), we set *QTR(n)* = 1; *QTR(n)* = 0, otherwise.[27] For the second reversal speed measure$\left(-\rho_1^{ST}\right)$, we similarly partition the observations into four quartiles and create four indicator variables corresponding to them. For example, we set *QTR(1)* = 1, for observations in the highest quartile, and 0 otherwise. Observe that these classifications provide ordinal measures of reversal speed. That is, for both reversal speed measures, the reversal is fastest for observations classified as *QTR(1)* = 1 and least quickly for observations classified as *QTR(4)* = 1. Thus, we anticipate more negative parameter estimates on interactions (*NOA_WC* * *QTR(n)*) for observations classified as having a higher reversal speed.

Table 5 reports results for specifications that use this approach. Estimates on the interactions (*NOA_WC* * *QTR(n)*) indicate the constraint on earnings management for observations classified as *QTR(n)*. Control variables are included in these specifications, but only the main effect and interaction effect for non-current accruals *NOA_NC* are tabulated. Statistical tests of restrictions of parameters on the four interaction terms, *NOA_WC* * *QTR(n)*, are displayed at the end of Table 5.

Estimates on the interactions in first column are –0.399, –0.157, –0.312, –0.048. Statistical tests at the bottom of the table indicate that the difference between estimate –0.399 on (*NOA_WC* * *QTR(1)*), and the estimate –0.048 on (*NOA_WC* * *QTR(4)*) is statistically significant at conventional levels (two-tailed $p < 0.01$). Though the magnitude of the estimate –0.157 on (*NOA_WC* * *QTR(2)*) is less than the estimate –0.312 on (*NOA_WC* * *QTR(3)*), the difference is not statistically significant. Thus, the evidence indicates a generally decreasing trend in the magnitudes of parameter estimates on the interactions as we move from *QTR(1)* to *QTR(4)*.

---

[27] From Table 1 (column 3), we know that, for 43 percent of our firm-quarters, current discretionary accruals reverse in the next quarter, i.e., $n^{ST} = 1$. This means *QTR(1)* = 1 for 43 percent of firm-quarters. Similarly, *QTR(2)* = 1 for 29 percent; *QTR(3)* = 1 for 21 percent; and *QTR(4)* = 1 for 7 percent of firm-quarters.

TABLE 5

Results for Reversal Speed Categories

Panel A: Explanatory Variables

	Parameter Estimates (z-statistics in parentheses)	
	Reversal Speed Proxy = 1/n (1)	Reversal Speed Proxy = $(-\rho_1)$ (2)
Intercept	−1.364***	−1.384***
	(−10.313)	(−11.291)
$NOA\_WC_{t-1} * QTR(1)$	−0.399***	−0.428***
	(−6.620)	(−5.392)
$NOA\_WC_{t-1} * QTR(2)$	−0.157**	−0.432***
	(−2.098)	(−4.892)
$NOA\_WC_{t-1} * QTR(3)$	−0.312***	−0.169**
	(−3.814)	(−2.415)
$NOA\_WC_{t-1} * QTR(4)$	−0.048	−0.153**
	(−0.373)	(−1.977)
$QTR(1)$	−0.038	−0.100*
	(−0.513)	(−1.794)
$QTR(2)$	−0.084	−0.086
	(−1.116)	(−1.605)
$QTR(3)$	−0.063	−0.040
	(−0.794)	(−0.799)
NOA_NC_{t-1}	−0.031*	−0.038***
	(−1.849)	(−2.976)
$NOA\_NC_{t-1} * speed^{LT}_{t-1}$	−0.075	0.064
	(−0.853)	(1.096)
$speed^{LT}_{t-1}$	−0.115	0.164*
	(−1.153)	(1.676)
Control Variables	Included	Included
Likelihood ratio	3,017.563	3,030.806

Panel B: Wald Tests for Selected Parameter Restrictions

Restrictions	Reversal Speed Proxy =1/n (1) Wald χ^2	Reversal Speed Proxy = $(-\rho_1)$ (2) Wald χ^2
1. $NOA\_WC * QTR(1) = NOA\_WC * QTR(2)$	6.720***	0.002
2. $NOA\_WC * QTR(2) = NOA\_WC * QTR(3)$	2.024	5.690**
3. $NOA\_WC * QTR(3) = NOA\_WC * QTR(4)$	3.161*	0.025
4. $NOA\_WC * QTR(1) = NOA\_WC * QTR(4)$	6.144***	6.250***
5. $NOA\_WC * QTR(1) = NOA\_WC * QTR(2) = NOA\_WC * QTR(3) = NOA\_WC * QTR(4) = 0$	10.165***	11.942***

*, **, *** Significance at the 0.10, 0.05, and 0.01 levels (two-tailed), respectively.

Entries are parameter estimates from logit specifications of *MBE*, the probability of meeting or beating the most recent quarterly consensus analyst forecast. Four indicator variables, *QTR(1)* to *QTR(4)*; indicate the four levels of our estimated reversal speed. In Column (1), the four levels of the reversal speed are based on our estimate of n.

(continued on next page)

TABLE 5 (continued)

For example, if current discretionary accruals initiated in quarter t reverse in quarter t+2, then *QTR(2)* = 1, 0 otherwise. In Column (2), four indicator variables are identified based on four quartiles of $(-\rho_1)$, where *QTR(1)* = 1 when $(-\rho_1)$ is in the highest quartile, 0 otherwise. Note that, for both classifications, *QTR(1)* to *QTR(4)* decreases in reversal speed. Other variables are defined in Table 2. z-statistics (in parentheses) are based on standard errors clustering on firm. Estimates for control and fiscal year indicator variables are not reported.

In Column 2 of Table 5, we report results for quartiles of ρ_1, the first-order autocorrelation of *NDA*. Estimates on interactions are −0.428, −0.432, −0.169, −0.153, respectively, as reversal speed declines from *QTR(1)* to *QTR(4)*. Again, the difference between estimate −0.428 on (*NOA_WC* * *QTR(1)*) and the estimate −0.153 on (*NOA_WC* * *QTR(4)*) is statistically significant (two-tailed $p < 0.01$).

These decreasing trends and related statistical comparisons support the hypothesis that the constraint on earnings management from the reversal of prior manipulation of current net operating assets (measured empirically as *NOA_WC*) varies directly with the speed of discretionary accrual reversal. Moreover, statistical tests displayed at the end of Table 5 of the four interaction terms *NOA_WC* * *QTR(n)* considered as a group indicate that variation in the likelihood of meeting or exceeding quarterly forecasts explained by the interactions is statistically significant (Wald χ^2 = 10.165 and 11.942; $p < 0.01$).

VI. CONCLUDING REMARKS

In this study, we model the process of discretionary accrual reversal, which establishes a link between net operating assets (NOA) and future earnings management. We postulate that, for a specified amount of cumulative bias in NOA at the beginning of a period, reversal speed determines the extent of the constraint on earnings management in the current period. We then perform an empirical analysis in the context of Barton and Simko (2002). In contrast with Barton and Simko (2002), which also advances NOA as a determinant of earnings management, we condition expectations on the reversal speed of prior discretionary accruals.

Using 36,210 firm-quarter observations from 2,109 firms, we find, first, that negative associations between the beginning-of-the-quarter NOA and the likelihood of achieving earnings targets documented by Barton and Simko (2002) dissipates after 2003, when regulation and oversight of financial reporting increases substantially. The evidence supports conclusions advanced in Cohen et al. (2008) that recent events, and/or recently imposed regulatory actions, discourage the use of accruals to manage toward earnings objectives.

We then focus on the period from 1993 to 2003 to consider whether a given bias in NOA at the beginning of a quarter imposes a greater constraint on current quarter's earnings management when the reversal of prior discretionary accruals occurs immediately (e.g., in the next quarter) or occurs more slowly. Evidence supports a characterization where faster reversal of prior current (but not non-current) discretionary accruals imposes a greater constraint on management's ability to bias subsequent earnings in the same direction. More specifically, we show that a given amount of NOA overstatement imposes different degrees of constraint on subsequent earnings management, depending on the reversal speed of discretionary accruals. This insight informs both policy and practice to the extent that it promotes understanding of the consequences of earnings management.

REFERENCES

Abarbanell, J., and R. Lehavy. 2002. Differences in commercial database reported earnings: Implications for empirical research. Working paper, the University of North Carolina at Chapel Hill, NC.

Aiken, L., and S. West. 1996. *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks CA: Sage Publications.

An, A. B. 2008. Performing logistic regression on survey data with the new SURVEYLOGISTIC procedure. *Statistics and Data Analysis*: 258–27.

Baber, W., S. Chen, and S. Kang. 2006. Stock price reaction to evidence of earnings management: Implications for supplementary financial disclosure. *Review of Accounting Studies* 11 (1): 5–19.

Barton J., and P. Simko. 2002. The balance sheet as an earnings management constraint. *The Accounting Review* 77 (Supplement): 1–27.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating analysts' earnings forecasts. *Journal of Accounting and Economics* 33 (2): 173–204.

Beneish, M. 1997. Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16 (3): 271–309.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (1): 99–126.

Cohen, D., A. Dey, and T. Lys. 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods. *The Accounting Review* 83 (3): 757–787.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 35–59.

Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 3–42.

DeFond, M. 2002. Discussion of: The balance sheet as an earnings management constraint. *The Accounting Review* 77 (Supplement): 29–33.

DeFond, M., and C. Park. 2001. The reversal of abnormal accruals and the market valuation of earnings surprises. *The Accounting Review* 76 (3): 375–404.

Dye, R. 1988. Earnings management in an overlapping generations model. *Journal of Accounting Research* 26 (2): 195–235.

Fan, Y., A. Barua, W. Cready, and W. Thomas. 2010. Managing earnings using classification shifting: Evidence from quarterly special items. *The Accounting Review* 85 (4): 1303–1323.

Francis, J., and M. Smith. 2005. A reexamination of the persistence of accruals and cash flows. *Journal of Accounting Research* 43 (3): 413–451.

Guay, W., S. Kothari, and J. Watts. 1996. Evaluation of discretionary accrual models. *Journal of Accounting Research* 34 (Supplement): 83–105.

Hunt, A., S. Moyer, and T. Shevlin. 1996. Managing interacting accounting measure to meet multiple. objectives: A study of LIFO firms. *Journal of Accounting and Economics* 21 (3): 339–374.

Jones, J. 1991. The effect of foreign trade regulation on accounting choices, and production and investment decisions. *Journal of Accounting Research* 29 (2): 193–228.

Kang, S., and K. Sivaramakrishnan. 1995. Issues in testing earnings management and an instrumental variable approach. *Journal of Accounting Research* 33 (2): 353–367.

Kirschenheiter, M., and N. Melumad. 2002. Can big bath and earnings smoothing co-exist as equilibrium financial reporting strategies? *Journal of Accounting Research* 40 (3): 761–796.

Kothari S. P., A. Leone, and C. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

Norton, E., H. Wang, and C. Ai. 2004. Computing interaction effects and standard errors in logit and probit models. *The Stata Journal* 4 (2): 154–167.

Rangan, S. 1998. Earnings management and performance of seasoned equity offerings. *Journal of Financial Economics* 50 (1): 101–122.

Richardson, S., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437–485.

Trueman, B., and S. Titman. 1988. An explanation of accounting income smoothing. *Journal of Accounting Research* 26 (Supplement): 127–139.

Verrecchia, R. 1986. Managerial discretion in the choice among financial reporting alternatives. *Journal of Accounting and Economics* 8 (3): 175–195.

White, H. 2010. An accrual balance approach to assessing accrual quality. Working paper, University of Michigan.

APPENDIX A
PROOF OF THE PROPOSITION

Assumptions in Section II imply that the autocorrelation structure of (d_t) is: $Cov[d_t, d_{t-k}] = Cov[\varepsilon_t, \varepsilon_{t-k}] = \theta^k\sigma^2$, where $0 \leq \theta < 1$, and θ^k is the *k*th order autocorrelation of ε_t. Denote ρ_k as the *k*th order autocorrelation coefficient of NDA_t $(d_t - d_{t-n})$ estimated using the Jones model.

Proposition 1: If period *t* discretionary accruals (d_t) reverse fully in period $t+n$, where $n \geq 1$, then the minimum value of ρ_k is achieved when $k = n$.

Proof: The *k*th order autocorrelation coefficient (ρ_k) of *NDA* is:

$$\rho_k = \frac{Cov(d_t - d_{t-n},\ d_{t+k} - d_{t+k-n})}{Var(d_t - d_{t-n})} = \frac{2\theta^k - \theta^{|k-n|} - \theta^{k+n}}{2(1-\theta^n)}.$$

Next we show that ρ_k varies inversely with $k < n$ and varies directly with $k \geq n$.

$$\text{For } k \geq n,\ \rho_k = \frac{2\theta^k - \theta^{k-n} - \theta^{k+n}}{2(1-\theta^n)}, \text{and} \frac{\partial\rho_k}{\partial k} = -\theta^k(\ln\theta)\frac{(1-\theta^n)}{2\theta^n} > 0;$$

$$\text{For } k < n,\ \rho_k = \frac{2\theta^k - \theta^{n-k} - \theta^{k+n}}{2(1-\theta^n)}, \text{and} \frac{\partial\rho_k}{\partial k} = (\ln\theta)\frac{2\theta^k + \theta^n\left(\frac{1}{\theta^k} - \theta^k\right)}{2(1-\theta^n)} < 0.$$

Thus, ρ_i is at its minimum for $k = n$.

$$Min(\rho_k) = \rho_{k=n} = \frac{2\theta^n - 1 - \theta^{2n}}{2(1-\theta^n)} = -\frac{(1-\theta^n)}{2}.$$

Proposition 2: The first-order autocorrelation coefficient of NDA_t (ρ_1) varies inversely with the reversal speed of discretionary accruals (d_t).

Proof: The first-order autocorrelation coefficient (ρ_1) of *NDA* is:

$$\rho_1 = \frac{Cov(d_t - d_{t-n},\ d_{t+1} - d_{t+1-n})}{Var(d_t - d_{t-n})} = \frac{2\theta\sigma^2 - \theta^{n-1}\sigma^2 - \theta^{n+1}\sigma^2}{2(1-\theta^n)\sigma^2}$$
$$= \frac{(\theta - \theta^{n-1})}{2(1-\theta^n)} + \frac{\theta}{2}.$$

The first-order condition when $\theta > 0$, is:

$$\frac{\partial\rho_1}{\partial n} = -\frac{\ln\theta}{2} * \frac{(\theta^{n-1})(1-\theta^2)}{(1-\theta^n)^2} + \frac{1}{2} > 0.$$

Reversal speed varies inversely with n. Thus:

$$\frac{\partial\rho_1}{\partial(reversal\ speed)} < 0.$$

Notice that, if discretionary accruals are uncorrelated across time ($\theta = 0$), then $Cov[d_t, d_{t-k}] = 0$, for $k \geq 1$. Then:

$$\text{for } n = 1, \rho_1 = \frac{Cov(d_t - d_{t-1},\ d_{t+1} - d_t)}{Var(d_t - d_{t-1})} = -0.5;$$

$$\text{for } n > 1, \rho_1 = 0.$$

Thus, Proposition 2 applies trivially.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1213–1230

American Accounting Association
DOI: 10.2308/accr-10038

Spillover Effects in Subjective Performance Evaluation: Bias and the Asymmetric Influence of Controllability

Jasmijn C. Bol
Steven D. Smith
University of Illinois at Urbana–Champaign

ABSTRACT: We examine how subjective performance evaluations are influenced by the level and controllability of an accompanying measure of a separate performance dimension. In our experiment, supervisors evaluate the office administration performance of a hypothetical subordinate. We find that supervisors' subjective evaluations are directionally influenced by an accompanying objective measure of sales performance, even after excluding participants who perceive informativeness across measures. Consistent with concerns for fairness and motivation, we also find an asymmetric uncontrollability effect—supervisors' evaluations are higher when an uncontrollable factor decreases the subordinate's sales (i.e., they compensate for bad luck), but are not lower when the uncontrollable factor increases the subordinate's sales (i.e., they do not punish for good luck). This evidence suggests that supervisors use discretion provided to evaluate performance on one task to adjust for perceived deficiencies in the evaluation of performance on other tasks. Our study integrates theories of cognitive bias and motivation, highlighting the need to consider the potentially interactive effects of different performance measures in multi-task settings.

Keywords: *subjective performance evaluation; performance measurement; cognitive bias; controllability; fairness.*

Data Availability: *Contact the authors.*

We appreciate helpful comments and suggestions from two reviewers, Ranjani Krishnan (editor), Steve Kachelmeier (senior editor), Alexander Brüggen, Margaret Christ, Brooke Elliott, Anne Farrell, Steve Glover, Gary Hecht, Kevin Jackson, Ken Koga, Mark Peecher, Mike Sandretto, Ira Solomon, Hun-Tong Tan, Bill Tayler, Wim van der Stede, Alan Webb, Jeff Wilks, David Wood, Mark Zimbelman, and workshop participants at Brigham Young University, Emory University, the University of Illinois, Nanyang Business School, and the 2009 AAA Annual and 2010 MAS Midyear Meetings. We also thank Rachel Birkey, Ben Ger, and Andrew Gordon for valuable research assistance. This research was funded by the Campus Research Board of the University of Illinois at Urbana–Champaign.

Editor's note: Accepted by Ranjani Krishnan.

Submitted: June 2009
Accepted: December 2010
Published Online: April 2011

I. INTRODUCTION

Organizational incentive systems often allow managerial discretion in the evaluation of employee performance (Murphy and Oyer 2003). Subjective performance evaluation allows managers to use noncontractible information to assess actions and efforts that objective measures (such as those produced by the accounting system) are not able to capture, creating a more complete depiction of employee performance (Bol 2008).[1] Subjectivity can therefore be useful in reducing risk to employees and improving the incentive alignment of the firm's performance measurement system (Baker et al. 1994; Bushman et al. 1996; Hayes and Schaefer 2000).[2]

Many performance measurement systems include both objective measures and subjective evaluations (Prendergast 1999; Gibbs et al. 2004). While designed to capture separate dimensions of employee performance, these different measurement types can influence each other. For example, theories from psychology and organizational behavior suggest that subjective judgments can be unduly influenced by an individual's knowledge of other, unrelated information (Nisbett et al. 1981; Bond et al. 2007). Such an influence could limit the complementary role of subjective evaluation in improving the overall informativeness of the measures used to evaluate performance.

We examine how supervisors' subjective performance evaluations are affected by the level and controllability of an objective measure of a separate aspect of performance. We analyze a two-dimensional employment setting in which an employee's performance on one dimension (i.e., task) is measured objectively, while performance on the other dimension is evaluated subjectively by the supervisor. In many such settings, the objective performance measure is known by the supervisor before s/he subjectively evaluates the employee's performance on the other dimension (Huber et al. 1987; Bommer et al. 1995). Our first research question examines whether the level of the objective measure has a directional impact on the supervisor's subjective evaluation. Specifically, we examine whether, consistent with cognitive distortion, supervisors bias their subjective evaluations of performance on one dimension to be consistent with an objective measure of performance on a separate and unrelated dimension.[3]

Our second research question considers how such a spillover effect differs when the controllability of the objective performance measure is relatively low. Uncontrollability introduces noise and error into the performance measurement system (Feltham and Xie 1994), and has been shown to affect attribution judgments (Tan and Lipe 1997). Employees are likely to perceive uncontrollable performance measures to be unfair when the uncontrollability reduces measured performance. Perceptions of unfairness in compensation lead to reduced job satisfaction and motivation (Cohen-Charash and Spector 2001; Colquitt et al. 2001). We examine whether this expectation is reflected in supervisors' use of their discretion, even when that discretion is provided for the evaluation of a separate task. Specifically, we examine whether supervisors use the discretion in their subjective evaluations to adjust for (or offset) the effects of an uncontrollable objective measure. Consistent with concerns for fairness and employee motivation, we predict that supervisors will make such an adjustment when the uncontrollable factor reduces measured performance (i.e., the employee suffers from bad luck) but not when it increases measured performance (i.e., the employee benefits from good luck).

---

[1] We define subjectivity as judgment based on personal impressions and opinions, which are not verifiable in a court of law and are therefore noncontractible.

[2] Subjectivity in performance evaluation can be introduced by the use of subjective performance measures, by allowing for *ex post* flexibility in the weighting of objective performance measures, or by allowing for *ex post* discretional adjustment based on factors other than the performance measures specified *ex ante* (Bol 2008). Our focus is on the first of these mechanisms.

[3] It is possible that supervisors' subjective evaluations may be influenced by other *subjective* measures as well as by objective measures. We specifically examine the effect of *objective* performance measures because such measures are likely to be known at the time the subjective evaluation is performed, and because of our interest in the interactive effect of controllability, which is most likely to be a feature of objective performance measures.

In our experiment, experienced supervisors employed by a large state university participate as evaluators in a hypothetical case setting. Participants assume the role of a regional director with supervisory authority over district managers who have both sales- and office administration-related duties. Objective individual sales information is given, after which participants are charged with subjectively evaluating the office administration performance of one district manager based on personal notes and staff interview responses, which are held constant in all conditions. The experiment employs a $2 \times 2 + 1$ between-subjects design. In four treatment conditions, we manipulate the level of the manager's individual sales at two levels by varying the objective sales score (high and low), and we manipulate the controllability of the individual sales measure at two levels by varying whether significant but uncontrollable events impacted the manager's sales during the period. A fifth (control) condition includes only the information about the manager's office administration performance.

We find that supervisors' subjective evaluations of the manager's office administration performance are significantly higher (lower) when the objective level of the manager's individual sales measure is relatively high (low). By excluding those participants who perceived the individual sales measure to be informative about the manager's office administration performance in our primary analysis, we provide evidence of cognitive distortion of performance information as a result of exposure to the objective measure.

We also find that the effect of the objective measure depends on its controllability. Specifically, participants use their subjective evaluations to adjust for the impact of uncontrollable events on the sales measure. Consistent with theory, this effect is asymmetric—participants' evaluations are higher (relative to the high controllability condition) when the uncontrollable factor leads to an unfavorable outcome, but are no lower when the uncontrollable factor leads to a favorable outcome. Thus, we find that participants use their discretion to compensate for bad luck, but not to punish for good luck.

This study contributes to the accounting and management literatures on performance evaluation. Prior research on bias in performance evaluation has found that supervisors' subjective evaluations of current performance may be directionally influenced by prior performance information (Murphy et al. 1985; Huber et al. 1987; Kravitz and Balzer 1992) or information from different sources (Blakely 1993; Murphy and Cleveland 1995; Bono and Colbert 2005). Our study contributes to this line of research by showing that cognitive distortion can cause spillover from the objective evaluation of one dimension of performance to the subjective evaluation of performance on a separate and unrelated dimension.

Prior research has also found evidence that controllability concerns can affect evaluators' judgments in single-dimension performance-evaluation settings. For example, Tan and Lipe (1997) find that controllability moderates outcome effects in evaluators' judgments (also see Brown and Solomon 1987). Our study reveals a very different controllability effect. Specifically, our results suggest that supervisors can use discretion that is provided for the evaluation of performance on a particular dimension to "correct" perceived deficiencies in other areas of the evaluation system. While the use of "borrowed" discretion can yield motivational benefits for employees, the intended benefits of that discretion (i.e., accurate reflection of performance on a specific dimension) can be reduced. Overall, our evidence indicates that the known level of an objective performance measure compromises the effectiveness of subjective performance evaluation for completing the picture of employee performance. Thus, our results have implications for the weight, timing, and nature of subjectivity in compensation contracting, suggesting that consideration of different measurement types in isolation can lead to incomplete conclusions about the optimal design of performance measurement systems.

An asymmetric pattern of compensation payouts, in which managers are sheltered from downside risk but are allowed the advantages of upside risk, has been demonstrated in accounting research (Gaver and Gaver 1998; Bertrand and Mullainathan 2001). While our study shows a similar pattern, it differs from this research by demonstrating the phenomenon of borrowed discretion, and by illustrating this effect in a performance-evaluation setting without the possible explanations of personal relationships, managerial entrenchment, or other issues related to corporate governance.

Our study also demonstrates that particular features of accounting information (in this case, the controllability of an objective performance measure) can produce differences in the spillover effects it induces in decision-makers' judgments. Predominantly, the effect of the objective performance measure is driven by its directional content, and the spillover effect takes the form of assimilation. However, when the objective performance measure is reduced by an uncontrollable factor, fairness and motivational concerns are triggered, inducing a corrective adjustment that offsets the directional bias. Thus, our study integrates theories of cognitive distortion and fairness, providing evidence that the interaction of objective and subjective performance measures depends on the fundamental qualities of the objective measure.

Section II presents background and hypotheses. Section III describes the experimental methodology, Section IV presents the results, and Section V concludes.

II. BACKGROUND AND HYPOTHESES

Prior Literature

Prior research has explored potential drawbacks of the use of subjective performance evaluation. These studies find that the role of human judgment, along with the inability to verify outcomes, can give rise to a number of problems that may limit the informativeness and incentive-strengthening role of subjective measures (Prendergast and Topel 1993, 1996; Bol 2008, 2011). Much of the research in this area has focused on supervisors' personal incentives to distort evaluations. For example, Bol (2011) finds that supervisors' tendency toward centrality and leniency in their evaluations is associated with incentive-related factors. Executives in Longenecker et al. (1987) confirm that they often prioritize incentive-consistent goals (e.g., higher raises for one's subordinates, improved motivation and encouragement, and signaling to poor performers) over accuracy in their evaluations.

A related literature has examined how subjective performance evaluations are influenced by cognitive limitations. Research on context effects examines how assimilation or contrast arises when evaluators have knowledge of an employee's performance in a prior period when evaluating his/her current-period performance (Murphy et al. 1985; Kravitz and Balzer 1992). Research on the halo effect finds that supervisors attend to general impressions when evaluating subordinates (Nisbett and Wilson 1977; Fox et al. 1983; Tan and Jamal 2001). Research in specific accounting domains has also explored the role of cognitive limitations in performance evaluation. Lipe and Salterio (2000) find that supervisors with discretion in the weighting of different performance measures underweight unique measures in favor of common measures. Ittner et al. (2003) provide evidence that supervisors place greater weight on outcome/results measures than on input/driver measures, and that they place more weight on objective or quantitative measures than on subjective or qualitative measures. Bailey et al. (2011) find that managers' reliance on the anchoring and adjustment heuristic prevents the full incorporation of non-contracted (i.e., subjective) information in the allocation of a bonus pool.

These studies suggest that cognitive limitations can preclude the full benefits of introducing subjectivity into compensation contracts. The purpose of our study is to examine the potentially interactive effects of objective and subjective performance measures. Specifically, we examine whether cognitive limitations preclude what should be independent assessments of separate and unrelated performance dimensions, and how the controllability of the objective measure influences this interaction.

Hypothesis Development

We consider a multi-task setting in which an employee's performance on one task is evaluated objectively, while performance on a separate task is evaluated subjectively.[4] In many such settings, the objective performance measure for a period is known before the subjective performance evaluation takes place (Huber et al. 1987; Bommer et al. 1995). For example, a supervisor is likely to observe an employee's individual sales or production performance in a period before conducting a formal evaluation of that employee's performance on more qualitative dimensions such as leadership, teamwork, creativity, communication, and attitude.

Because subjective performance evaluations cannot be verified by outside parties, supervisors could be influenced by their own preferences when evaluating performance (Prendergast and Topel 1993). We argue, however, that even in the absence of personal incentives to bias evaluations, supervisors are likely to be influenced by the observed level of performance on one task when subjectively evaluating an employee's performance on a separate task. We also argue that the form of an objective measure's "spillover" will depend on its controllability, or the degree to which an employee can affect the measure by altering his/her action inputs (Antle and Demski 1988).

Cognitive Distortion

The supervisor's knowledge of an employee's objective performance on one dimension is likely to affect his/her subjective evaluation of performance on a separate dimension by inducing *cognitive information distortion*. Specifically, supervisors will unintentionally bias their subjective evaluations to be consistent with the known level of performance on the objective measure. Prior research shows that individuals tend to unknowingly process information in a way that favors a previously held belief, desired outcome, or developing preference (Wilks 2002; Bond et al. 2007).

Bond et al. (2007) argue that, in the evaluation of a single option, information observed early evokes an "evaluative disposition," and that the subjective encoding, weighting, and interpretation of subsequent information is influenced to be coherent with that disposition. In the case of multi-task performance evaluation, this occurs when the known level of performance on one dimension causes the supervisor to unknowingly process the available information about performance on a separate dimension in the same direction. That is, a supervisor who observes a relatively high level of performance on one dimension is likely to interpret information about the subordinate's performance on a separate dimension more favorably than a supervisor who observes a relatively low level of performance on the first dimension.[5]

---

[4] We assume a system in which the supervisor is not allowed any other forms of discretion (for example, subjective weightings or subjective adjustment to objective measures). In other words, the only discretion resides in the supervisor's evaluation of performance on one of the two tasks.

[5] Cognitive distortion is related to but conceptually different from other context effects, such as anchoring and adjustment and confirmation bias (Bond et al. 2007). For example, there is an emerging consensus that the anchoring and adjustment heuristic arises due to biased accessibility of information from memory (Gilovich 1991; Baron 1995; Chapman and Johnson 1999; Mussweiler and Strack 1999), while confirmation bias reflects a bias in the search for information from the environment (Frey 1986; Klayman and Ha 1987). Because our study relates to the subjective interpretation of information (rather than how it is obtained), cognitive distortion (a processing-related phenomenon) is the most fitting explanation for the mechanism underlying the directional bias we predict.

As mentioned, objective measures of performance are likely to be known by supervisors when they conduct subjective evaluations of subordinates' performance. The literature cited above suggests that this knowledge will generate a directional bias that affects subsequent judgments. Thus, we predict that, consistent with cognitive distortion, supervisors' subjective performance evaluations will be influenced in the direction of their subordinates' objective performance levels.

> **H1:** Supervisors' subjective performance evaluations will be higher (lower) when the employee's performance on an unrelated objective measure is higher (lower).

It is important to note that if the objective level of performance on one task is seen by the supervisor as informative about (i.e., related to) the employee's performance on the other task, then the supervisor may *intentionally* incorporate information from the objective measure into his/her subjective evaluation. Economics-based models of performance measurement suggest that any (costless) measure that provides incremental information about employee performance should be included in the evaluation (Holmstrom 1979; Banker and Datar 1989). This reasoning can be intuitively extended to the use of information from a measure of performance on one task to the (subjective) evaluation of performance on another task. However, our interest in this study is in spillover effects arising from cognitive processes independent of supervisors' perceptions of informativeness. Thus, our experiment creates a setting in which the two performance measures are unrelated, and in our main analysis we exclude participants who nevertheless perceive informativeness across the two measures in order to provide the most conservative test of our theory.[6]

Controllability

In this section we examine how the controllability of an objective measure can influence the spillover effect it induces in supervisors' judgments. Controllable performance measures are preferred from a contracting perspective because uncontrollability reduces the informativeness of a measure about employee actions and effort, and because agents need to be compensated for the additional risk of uncontrollable factors (Holmstrom 1979). In practice, however, measures that are perfectly controllable are extremely rare. Indeed, noise as a common feature of objective performance measures is among the primary reasons for the use of subjective performance evaluation (Baker et al. 1994).

When an objective measure is relatively low in controllability, its spillover effect will take a different form. We predict that uncontrollability in the objective performance measure will trigger the supervisor's consideration of the measure's quality and its effect on the employee's overall evaluation. Specifically, supervisors have incentives to consider the fairness of performance evaluations, not only due simply to their individual preferences for justice (Fehr and Fischbacher 2002), but also (and perhaps more importantly) because employee perceptions of unfairness in evaluation and compensation have a negative effect on satisfaction and motivation (Greenberg 1987; Cohen-Charash and Spector 2001; Colquitt et al. 2001). Bol et al. (2010) find that supervisors adjust performance targets to address fairness concerns arising within compensation contracts.

Because of the importance of employee satisfaction and motivation for productivity, we predict that supervisors will use the discretion available to them to "correct" or adjust for an uncontrollable factor, even if that discretion is provided for a separate purpose. However, an adjustment for an uncontrollable factor is likely to depend on its directional impact. When an uncontrollable factor causes measured performance to be lower than it would likely have been otherwise, the employee is likely to feel that the measure does not capture his/her real efforts or contribution to the firm, and that this may result in lower compensation. As a result, s/he will consider the measure unfair and suffer reduced satisfaction and motivation. In order to protect the employee from this downside risk and to avoid the negative effects on employee motivation, the supervisor can exercise the discretion *in his/her subjective evaluation* by adjusting upward (i.e., giving a relatively higher evaluation) to countervail the relatively low level of the uncontrollable measure. In this case, the supervisor's subjective evaluation is still biased downward because of exposure to the low objective measure, but concern for the employee's motivation offsets the bias, leading the supervisor to adjust upward to ensure an *overall* evaluation that the employee will perceive as fair.

---

[6] Despite efforts to create orthogonal performance dimensions, we are unable to *ex ante* eliminate a conscious influence based on perceived informativeness as a partial explanation of our results. In many settings, tasks are bundled together for a single employee because of presumed complementarities in the performance of those tasks (Milgrom and Roberts 1992). Thus, it would be difficult to argue complete unrelatedness for a multi-task experimental setting that is designed to be realistic.

In contrast, when an uncontrollable factor causes measured performance to be higher than it would have been otherwise, the employee is unlikely to perceive unfairness or to feel reduced motivation.[7] In this case, the supervisor's use of discretion to "correct" the discrepancy between measured and actual performance (i.e., by giving a relatively lower evaluation) would likely be considered an inappropriate and unfair departure from the prescribed evaluation procedure. Thus, while the supervisor is likely to recognize that the measure fails to capture the employee's true contribution, s/he has no incentive to correct it, and the directional bias predicted in H1 is not expected to be offset.

In summary, we predict that supervisors' use of discretion to adjust for uncontrollability in a separate performance measure follows an asymmetric pattern. When an uncontrollable factor leads to an unfavorable outcome, supervisors' consideration of the employees' fairness perceptions and motivation offsets the directional bias, and we predict higher evaluations relative to the high controllability condition. However, when an uncontrollable factor leads to a favorable outcome, supervisors will not use their discretion to correct, and we predict no difference in subjective evaluations relative to the high controllability condition.

H2a: When the employee's performance on an unrelated objective measure is relatively low, the supervisor's subjective performance evaluation will be higher (lower) when the objective measure is (is not) reduced by an uncontrollable factor.

H2b (null hypothesis): When the employee's performance on an unrelated objective measure is relatively high, the supervisor's subjective performance evaluation will not be affected by whether the objective measure is increased by an uncontrollable factor.

III. METHOD

Participants completed a performance evaluation task in a hypothetical case setting, assuming the role of a regional director for an industrial pipe and fitting company. The regional director's job was described as overseeing a number of district managers who were responsible for both individual sales and district office administration. The experimental task was to perform a subjective evaluation of the office administration performance of one of the district managers, while the manager's score on an objective individual sales measure was known.

---

[7] This is consistent with the well-established principle that unfairness is felt more (less) acutely when distributed rewards are less (more) than what is felt to be just (Jasso 1978). Similarly, prior research in organizational behavior finds that consideration of the fairness of the evaluation procedure is triggered when the outcome is low, but not when it is high (Cropanzano and Folger 1989; Cropanzano and Konovsky 1995).

Participants

Supervisors (n = 216) at a large, public university participated in the experiment. Participants were targeted through a university database search for non-academic employees in supervisory positions. They received a letter by mail, inviting them to participate. A follow-up email invitation was sent about two weeks later. As an incentive to complete the study, all participants' names were entered into 20 random drawings, each worth $200.

Participants reported an average of 30 years of total work experience and 13 years of supervisory experience. In addition, participants supervised an average of nine employees in their current positions. Participants reported an average of ten years of performance-evaluation experience, and 60 percent reported having either formal or informal performance-evaluation duties for an average of six subordinate employees in their current positions. Finally, participants reported that about 50 percent of their performance-evaluation duties involve subjective evaluation (versus objective measurement).[8]

Experimental Design

The experiment employs a 2 × 2 + 1 between-subjects design. We manipulate the district manager's score on the sales measure at two levels (high and low), and we manipulate the controllability of the objective sales measure at two levels (high and low). We also add a control condition in which no objective sales information is provided.

Independent Variables

In the four treatment conditions, the objective measure used in the performance-evaluation system is the district manager's individual sales during the most recent six-month period. The measure is scored on a 0–10 scale, with each score connected to a particular range of sales.

The level of the target district manager's performance on the sales dimension is manipulated by varying his individual sales score. Specifically, the district manager's individual sales score is stated as 2 (9) in the low (high) sales condition.

Controllability is manipulated by varying the degree to which outside factors influenced the manager's individual sales. In the high controllability condition, participants are told that market conditions were stable during the period, and that there were no unforeseen or uncontrollable events that impacted the manager's sales. In the low controllability condition, participants are told that an uncontrollable event—specifically, either the opening (when the sales measure was high) or closing (when the sales measure was low) of large facilities by major customers in the manager's district—had a major impact on the manager's sales during the period.

Dependent Variables

Our primary dependent variable is the subjective evaluation of the district manager's office administration performance. In order for participants to make this evaluation, the case information included a collection of personal notes and interview responses from the manager's office staff. This material was the same in all conditions. After reading the case information, participants provided a performance-evaluation score, on a scale from 0 to 10, reflecting their assessment of the district manager's office administration performance. They also answered a number of post-experimental questions.

---

[8] We include participants' gender, years of work and supervisor experience, and number of evaluated subordinates as covariates in our statistical analyses. Results show that none of these variables interacts with the effects of interest.

Procedure

Participants completed an online experiment at their convenience by entering a URL into their web browsers. The survey program randomly assigned each participant to one of the five conditions (four treatment conditions and one control condition).

The case information asked each participant to assume a hypothetical situation in which he/she worked as a regional director for an industrial pipe and fitting company. A compressed organizational chart showed that the regional director had supervisory authority over ten district managers, who in turn supervised a number of sales associates and office staff. District managers' responsibilities and the associated performance measures were described.

The case informed participants that one of the regional director's duties was to conduct semiannual performance evaluations for each district manager under his/her supervision. The performance-evaluation system was described as comprising two components: an objective score based on individual sales and a subjective evaluation of the district manager's office administration performance. The instructions noted that the two components were independent because the job functions they captured were independent (i.e., performance on one job function was unrelated to performance on the other). Each component was scored on a 0–10 scale, and the two were averaged together (i.e., equally weighted) to arrive at the overall evaluation score. Participants were told that their task was to conduct the performance evaluation for one of the district managers. For the objective measure, a table showing the district managers' individual sales and their associated scores on the 0–10 scale was provided. The target district manager's sales were manipulated to be either high (9) or low (2). The manipulated controllability information was also given at this point.

For the subjective evaluation, participants were told that they had made some personal notes about the district manager over the six-month period, and those were provided. In addition, participants were told that because they were unable to observe the district manager's office administration performance on a daily basis, they had conducted interviews with a few members of his office staff, and the interview responses were also provided. The personal notes and interview responses were the same for all participants, and were designed to portray an average level of office administration performance. In the control condition, participants were asked to provide a subjective office administration evaluation, but were not given any individual sales information.

After reading all case information, participants were asked to indicate a score, on a 0–10 scale, reflecting their assessment of the district manager's office administration performance. They were also asked to average the two performance-evaluation components to arrive at the overall evaluation, and to assess their confidence in their judgments. They were then asked a number of post-experimental processing and demographic questions.

IV. RESULTS

Manipulation Checks and Information Sufficiency

We included two manipulation check questions in the post-experimental questionnaire. Participants recalled that the target manager's individual sales score was higher in the high sales condition than in the low sales condition ($p < 0.01$). Participants also judged the controllability of the sales measures to be higher in the high controllability condition than in the low controllability condition ($p < 0.01$). Thus, our manipulations of the level and controllability of the manager's individual sales measure were successful.

We also asked participants to rate (on a six-point scale) their agreement with a statement that the personal notes and interview responses about the target manager allowed them to make a reasonable judgment about the manager's office administration performance. The mean of this measure was 4.75 (s.d. = 1.05), significantly higher than the scale midpoint of 3.5 ($p < 0.01$). We also asked participants to rate (on a seven-point scale) the confidence they felt in their evaluation judgments. The mean of this measure was 5.5 (s.d. = 1.1), significantly higher than the scale midpoint of 4.0 ($p < 0.01$). Together, these results suggest that participants felt they had sufficient information on the manager's office administration performance to make a reasonable assessment.

Table 1 shows the mean and standard deviation of office administration performance scores by experimental condition. Figure 1 displays the means of the high and low controllability conditions (at high and low sales levels), as well as the control condition. The figure shows the expected pattern of results; the overall mean evaluation is highest in the high sales condition (6.90), followed by the control condition (6.35) and the low sales condition (5.42), and the predicted asymmetric pattern of treatment means is apparent across the two controllability conditions. Preliminary analyses (untabulated) show that the mean of the control condition is significantly lower than the mean of the high sales condition ($t = 2.54$; $p = 0.01$) and significantly higher than the mean of the low sales condition ($t = 3.67$; $p < 0.01$). These results confirm that the mean in the control condition is not so extreme that floor or ceiling effects would prevent the hypothesized results, and they reveal a directional spillover effect occurring in both directions.

Hypothesis Tests

Cognitive Distortion

H1 predicts that participants' subjective office administration evaluations will be higher (lower) when the individual sales measure is higher (lower). In order to conduct the cleanest and most conservative test of our theory, we make the following adjustments. First, because H2a predicts that low controllability will work against a main effect of sales level, we test H1 by comparing the low and high sales conditions only under high controllability.

TABLE 1

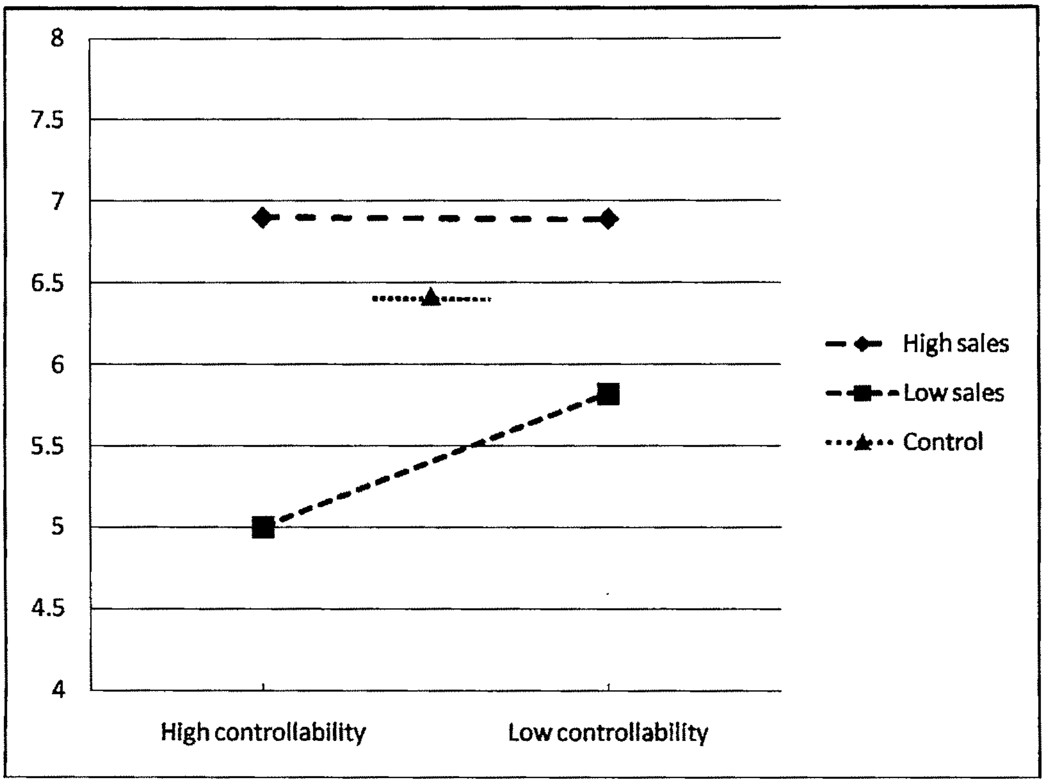
Descriptive Statistics: Least Square Means (Std. Errors) of Evaluation Scores by Sales Measure Level and Controllability

	High Controllability	**Low Controllability**	**Row Mean**
High sales	6.90 (1.14) n = 42	6.89 (0.95) n = 46	6.90 (1.04) n = 88
Low sales	5.00 (1.31) n = 42	5.82 (1.72) n = 44	5.42 (1.58) n = 86
Column mean	5.95 (1.55) n = 84	6.37 (1.47) n = 90	

Control Condition:
6.35
(1.21)
n = 42

Sales level is manipulated by varying whether the target sales manager achieved a relatively high or low sales score during the period (where each sales score was associated with a particular range of sales). Controllability is manipulated by varying whether an event outside the manager's control significantly affected his sales during the period. In the control condition no objective sales performance information was provided.

FIGURE 1
Mean Subjective Evaluations, High and Low Controllability, and Control Conditions



Second, as mentioned, it is possible that supervisors perceived that the sales measure contained information about the manager's office administration performance, and incorporated that information into their subjective evaluations. In order to isolate the role of cognitive distortion, we asked participants (using a Yes/No question in the post-experiment questionnaire) whether they believed the manager's individual sales score told them something about how well the manager performed as an office administrator. Within the high controllability condition, 65 percent of participants did not consider the sales measure to be informative for their subjective evaluations.[9] Our primary test of H1 includes only these participants.[10]

---

[9] Responses to this question varied across cells. Participants considered the objective measure to be significantly less informative when controllability was low (73 percent answering no) than when it was high (65 percent; $p < 0.01$) and considered the objective measure to be significantly less informative when the sales measure was high (77 percent) than when it was low (61 percent; $p < 0.01$).

[10] The relatively high proportion of participants who perceived informativeness across the two measures does not arise from a misunderstanding of the case materials. In the post-experimental questionnaire, these participants indicated that they understood the two task dimensions were independent of each other (mean agreement = 6.0 on a seven-point scale). It is likely that these participants merely did not believe that the two dimensions were completely unrelated. While they were well understood, the case instructions regarding independent dimensions are at odds with a deep-seated tendency to expect employee performance on multiple tasks to be correlated (Newcomb 1931; Jacobs and Kozlowski 1985). We thank a reviewer for pointing this out.

We test H1 with a one-way analysis of variance (ANOVA), comparing participants' subjective evaluations by sales level in the high controllability condition. The results, provided in Table 2, Panel A, indicate a significant effect of sales level ($F = 16.23$; $p < 0.01$), supporting H1. Not surprisingly, untabulated results also show a significant effect of sales level for those participants who did find the sales measure to be informative ($F = 91.52$; $p < 0.01$) and for the full sample of all high controllability participants ($F = 50.54$; $p < 0.01$). In sum, our results provide strong support for H1, suggesting that the effect of the sales measure on participants' subjective evaluations is driven, at least in part, by cognitive distortion of the case information.[11]

Controllability

H2a predicts that, when the employee's performance on an unrelated objective measure is relatively low, the supervisor's subjective performance evaluation will be higher (lower) when the objective measure is (is not) reduced by an uncontrollable factor. We test H2a with an ANOVA comparing the evaluations by controllability level in the low sales condition.[12] The results, provided in Table 2, Panel B, show a significant effect of controllability ($F = 6.15$; $p = 0.02$), supporting H2a.

H2b predicts that when the employee's performance on an unrelated objective measure is relatively high, the supervisor's subjective performance evaluation will not differ depending on whether the objective measure is increased by an uncontrollable factor. We test H2b with an ANOVA comparing the evaluations by controllability level in the high sales condition. The results, provided in Table 2, Panel C, show that the mean evaluation in the high sales, low controllability condition (6.89) is not significantly different from the mean in the high sales, high controllability condition (6.90) ($F = 0.00$; $p = 0.95$), failing to reject the null hypothesis H2b.

To address the possibility that the difference in controllability effects arose because participants perceived the effect of the uncontrollable factor on the quality of the sales measure differently across sales levels, we asked participants to rate (on a six-point scale) the extent to which the sales measure accurately reflected the manager's individual sales efforts. The mean ratings in the high sales, low controllability condition (mean = 3.59; s.d. = 1.27) and in the low sales, low controllability condition (mean = 3.25; s.d. = 1.57) are not statistically different ($p = 0.27$), yet both means are significantly lower than the mean rating in the high controllability conditions (mean = 4.88; s.d. = 1.22) ($p < 0.01$). Thus, participants perceived the sales measure to be similarly problematic in both low controllability conditions, yet they only adjusted for the measure when the uncontrollable factor reduced sales.

Supplemental Analyses

In order to gain a better understanding of participants' thought processes in the low controllability conditions, we asked participants to rate (on a seven-point scale) the extent to which

---

[11] Another potential explanation is that participants merely biased their subjective evaluations toward the objective measure levels to avoid the effort of processing the office administration information when conducting their evaluations. However, we would expect participants who took this approach to have poorer recall of the case information and/or to have spent less time completing the materials. Two separate analyses suggest that this does not explain the results. First, the post-experiment questionnaire included two quiz-type questions designed to confirm that participants had read the case materials carefully. Eighty-seven percent of participants correctly answered both questions, and results are identical when those who did not are excluded from the analyses. Second, we conduct an analysis of covariance using the amount of time participants spent on the experiment (used here as a proxy for effort) as a covariate. Results (untabulated) show that the amount of time taken does not interact with the effect of the sales level.

[12] We use the full sample of participants in the low controllability conditions for our tests of H2a and H2b because our primary interest is in participants' reactions to uncontrollability in the objective measure *regardless* of whether they perceive the measure to be informative for their subjective evaluations.

TABLE 2

Subjective Evaluation Results

Panel A: Analysis of Variance for Sales Level, High Controllability Condition

Source	df	Mean Square	F-statistic	p-value
Sales level	1	21.89	16.23	< 0.01
Error	52	1.35		

Panel B: Analysis of Variance for Controllability, Low Sales Condition

Source	df	Mean Square	F-statistic	p-value
Sales level	1	14.38	6.15	0.02
Error	84	2.34		

Panel C: Analysis of Variance for Controllability, High Sales Condition

Source	df	Mean Square	F-statistic	p-value
Sales level	1	0.01	0.00	0.95
Error	86	1.09		

See Table 1 for descriptions of sales level and controllability.

they considered what the manager's overall evaluation score would be when determining the office administration performance score. Within the low controllability conditions, 40 percent of participants provided ratings of 4 or higher, indicating that they did consider what the overall score would be (mean = 3.0).[13] To examine whether consideration of the overall score led participants to adjust for controllability in their evaluations, we regress the subjective evaluation score (*SCORE*) on the extent to which participants considered the overall score (*CONSIDER_OVERALL*). We control for participants' perceptions of the quality of the sales measure (*REFLECT_EFFORT*), as participants who believe the sales measure accurately reflects the manager's individual sales efforts are less likely to make a controllability adjustment. We run separate regressions for the low sales, low controllability and the high sales, low controllability conditions since adjustments in those two conditions would likely occur in opposite directions.

The results, presented in Table 3, show a significant positive relationship between consideration of the overall score and the subjective evaluation for participants in the low sales, low controllability condition (column 2). In contrast, there is no effect of consideration of the overall score for participants in the high sales, low controllability condition (column 3). These results support our theory regarding differences in participants' thought processes across the two low controllability conditions. While participants in both conditions considered what the overall evaluation would be, only when the manager's sales *suffered* from an uncontrollable factor did that consideration affect their subjective evaluations.

Participants' primary motivation for adjusting upward for controllability in the low sales, low controllability condition could have arisen from their own personal preferences for justice or from their concerns about the manager's satisfaction and subsequent motivation. That is, uncontrollability that reduces measured performance could cause dissatisfaction because the employee feels the evaluation is unfair and/or because it may reduce compensation. In an effort to explore adjustment

---

[13] The mean rating in the high controllability conditions is 2.9, with 32 percent providing ratings of 4 or higher.

TABLE 3

Regression of the Determinants of the Subjective Performance Score in the Low Controllability Condition

$$SCORE_i = \beta_0 + \beta_1 CONSIDER\_OVERALL_i + \beta_2 REFLECT\_EFFORT_i + \varepsilon_i$$

	SCORE[a]	
Variable	**Low Sales**	**High Sales**
CONSIDER_OVERALL	0.24*	0.03
	(1.91)	(0.45)
REFLECT_EFFORT	−0.30*	0.12
	(−1.84)	(1.09)
R^2	0.13	0.03

* Denotes statistical significance at the 10 percent level (two-tailed), with t-statistics in parentheses. n = 44 for the low controllability, low sales condition, and n = 46 for the low controllability, high sales condition.

[a] The intercept is included, but not separately reported.

Variable Definitions:
SCORE = participants' subjective evaluations of the manager's office administration performance;
CONSIDER_OVERALL = participants' ratings of whether they thought about what the manager's overall evaluation score would be when considering his office administration performance score (measured on a seven-point scale); and
REFLECT_EFFORT = participants' ratings of whether they believed that the manager's individual sales score accurately reflected his individual sales efforts (measured on a six-point scale).

motivations, we directly asked participants whether they used the flexibility (i.e., subjectivity) in their office administration evaluation to "correct" the manager's sales score. A follow-up question asked participants who indicated that they did correct what their primary motivation was. However, only two participants (4.5 percent) in the low sales, low controllability condition indicated that they made a corrective adjustment. This result suggests either that participants were unaware of their corrective adjustments or that they were unwilling to admit that they adjusted. As a result, we are unable to identify the primary motivations underlying participants' adjustments of the subjective evaluations in light of the uncontrollability of the sales measure. While we expect that all of the reasons discussed played a role in the adjustment effect we observe, a greater understanding of their relative importance is a potential avenue for future research.

V. CONCLUSION

We examine how objective measures of performance on one dimension affect supervisors' subjective evaluations of performance on a separate dimension. Our results indicate that supervisors distort information about employee performance to be directionally consistent with an objective measure of a separate task. However, when the objective measure is influenced by uncontrollable factors, we find that supervisors asymmetrically adjust for the resulting unfairness. Specifically, they use the discretion provided for the evaluation of a *separate* task to adjust upward to offset an uncontrollable and unfavorable outcome, but do not similarly adjust downward to offset an uncontrollable and favorable outcome.

Our study highlights the complex nature of subjective performance evaluation, which arises because the effect of discretion can take a number of different forms. Our results suggest that discretion can give rise to cognitive biases for supervisors in some situations, while allowing them to adjust for perceived measurement deficiencies in others. Importantly, we find these effects in a single-period setting that is free from personal relationships, incentives, or potential confrontations. These real-world features would likely make the effects we observe even stronger in many cases.

Our results have implications for the design of compensation contracts because they reflect a pattern of influence that has not been considered previously. There are many settings in which supervisors have discretion to subjectively evaluate performance that would be difficult to capture using objective measures. However, if supervisors are influenced by cognitive distortion, or if they instead use their discretion to overcome perceived deficiencies in the measurement of other tasks, their subjective evaluations are likely to be uninformative or even misleading. Such effects could have negative consequences in terms of compensation, promotion decisions, training, and motivation. Thus, contract designers who seek to capitalize on the complementary roles of objective and subjective performance evaluation should consider the effects we document here.

It is important to note that, while cognitive distortion could be seen as unambiguously detrimental, corrective adjustment may not necessarily be objectionable from an organizational perspective. Fairness perceptions play a significant role in the motivational effects of performance-evaluation systems (Cohen-Charash and Spector 2001; Colquitt et al. 2001). If supervisors are able to achieve a more fair evaluation outcome by using their discretion to overcome deficiencies in objective measures, then the benefits of increased employee motivation and satisfaction could outweigh the costs of uninformative subjective measures.

To the extent these effects are undesirable, however, a number of possible remedies warrant further examination. For example, it may be possible in some settings for supervisors to gather information and complete their subjective evaluations before observing objective measures of other performance dimensions. This timing adjustment would eliminate the risks of distortion and adjustment. Training evaluators to avoid these influences may also reduce the effects. Another possible approach is to introduce more subjectivity into the compensation contract. Controllability problems are primarily risk problems (i.e., they expose employees to risks beyond their control), and may be best addressed by allowing supervisors to subjectively adjust objective measures to account for uncontrollable events. The provision of additional, "targeted" discretion would prevent the need for supervisors to use discretion designed for one purpose (e.g., to evaluate performance on a particular dimension) to achieve other objectives (e.g., to correct for deficiencies in other areas of the evaluation system). However, greater subjectivity imposes greater costs, and the benefits of these and other interventions must be weighed against those costs.

The limited scope of our experimental setting presents a number of opportunities for future research. For example, all participants received the qualitative information for their subjective evaluations *after* observing the manager's sales score. Yet the generalizability of the distortion and adjustment results we observe may not be limited to this order of information presentation. Because the qualitative information is ambiguous and confronts the supervisor with a difficult decision under uncertainty, the objective measure (even if observed after the qualitative information is gathered) is still the first numerical value encountered, and could still influence the subjective evaluation. In addition, our experiment involved a one-period setting in which participants' predispositions could come from only one source (the current-period objective measure). However, in many settings supervisors have prior performance information and/or evaluations from other sources, all of which could provide competing beliefs about the manager. Future research could therefore examine the sensitivity of the effects we observe to the order in which information is presented and to the presence of other information that may affect the supervisor's predispositions toward the manager.

Participants in our study were not required to justify their evaluation judgments, yet accountability could potentially impact supervisors' susceptibility to these effects. Additionally, the information about the manager's office administration performance was limited in our experiment, and was designed to be relatively ambiguous (i.e., it contained both positive and negative signals). Variations in the amount and/or the directional strength of information a supervisor has about an employee's performance may influence the magnitude of distortion and adjustment effects. Similarly, the directional implication of the uncontrollable factor in our experiment was unambiguous (i.e., it clearly increased or clearly decreased the manager's individual sales). A more general uncontrollability would likely reduce the adjustment effects we observe because conclusions about the appropriate direction of any adjustment would be more difficult to draw. Finally, our experiment required participants to evaluate only a single employee. Yet many supervisors conduct evaluations for more than one employee, and the comparative element of multiple evaluations could affect the role of fairness and motivation. Future research could examine, for example, how supervisors consider issues such as fairness and controllability at the group level (rather than at the individual level) when using discretion in their subjective evaluations.

REFERENCES

Antle, R., and J. S. Demski. 1988. The controllability principle in responsibility accounting. *The Accounting Review* 63 (4): 700–718.

Bailey, W. J., G. Hecht, and K. L. Towry. 2011. Dividing the pie: The influence of managerial discretion extent on bonus pool allocation. *Contemporary Accounting Research* (forthcoming).

Baker, G. P., R. Gibbons, and K. J. Murphy. 1994. Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics* 109 (4): 1125–1156.

Banker, R. D., and S. M. Datar. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27 (1): 21–39.

Baron, J. 1995. Myside bias in thinking about abortion. *Thinking and Reasoning* 7: 221–235.

Bertrand, M., and S. Mullainathan. 2001. Are CEOs rewarded for luck? The ones without principals are. *The Quarterly Journal of Economics* 116 (3): 901–932.

Blakely, G. L. 1993. The effects of performance rating discrepancies on supervisors and subordinates. *Organizational Behavior and Human Decision Processes* 54: 57–80.

Bol, J. C. 2008. Subjectivity in compensation contracting. *Journal of Accounting Literature* 27: 1–32.

Bol, J. C. 2011. The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review* 86 (5) (forthcoming).

Bol, J. C., T. M. Keune, E. M. Matsumura, and J. Y. Shin. 2010. Supervisor discretion in target setting: An empirical investigation. *The Accounting Review* 85 (6): 1861–1886.

Bommer, W. H., J. L. Johnson, G. A. Rich, P. M. Podsakoff, and S. B. Mackenzie. 1995. On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology* 48: 587–605.

Bond, S. D., K. A. Carlson, M. G. Meloy, J. E. Russo, and R. J. Tanner. 2007. Information distortion in the evaluation of a single option. *Organizational Behavior and Human Decision Processes* 102: 240–254.

Bono, J. E., and A. E. Colbert. 2005. Understanding responses to multi-source feedback: The role of core self-evaluations. *Personnel Psychology* 58: 171–203.

Brown, C. E., and I. Solomon. 1987. Effects of outcome information on evaluations of managerial decisions. *The Accounting Review* 62 (3): 564–577.

Bushman, R. M., R. J. Indjejikian, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting and Economics* 21 (2): 161–193.

Chapman, G. B., and E. J. Johnson. 1999. Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes* 79 (2): 115–153.

Cohen-Charash, Y., and P. E. Spector. 2001. The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes* 86 (2): 278–321.

Colquitt, J. A., D. E. Conlon, M. J. Wesson, C. O. L. H. Porter, and K. Yee Ng. 2001. Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology* 86 (3): 425–445.

Cropanzano, R., and R. Folger. 1989. Referent cognitions and task decision autonomy: Beyond equity theory. *Journal of Applied Psychology* 74 (2): 293–299.

Cropanzano, R., and M. A. Konovsky. 1995. Resolving the justice dilemma by improving the outcomes: The case of employee drug screening. *Journal of Business and Psychology* 10 (2): 221–243.

Fehr, E., and U. Fischbacher. 2002. Why social preferences matter—The impact of non-selfish motives on competition, cooperation and incentives. *Economic Journal* 112 (478): 1–33.

Feltham, G. A., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69 (3): 429–453.

Fox, S., A. Bizman, and E. Herrmann. 1983. The halo effect: Is it a unitary concept? *Journal of Occupational Psychology* 56 (4): 289–296.

Frey, D. 1986. Recent research on selective exposure to information. In *Advances in Experimental Social Psychology*, Vol. 19, edited by L. Berkowitz, 41–80. New York, NY: Academic Press.

Gaver, J., and K. M. Gaver. 1998. The relation between nonrecurring accounting transactions and CEO cash compensation. *The Accounting Review* 73 (2): 235–253.

Gibbs, M., K. A. Merchant, W. A. Van der Stede, and M. E. Vargus. 2004. Determinants and effects of subjectivity in incentives. *The Accounting Review* 79 (2): 409–436.

Gilovich, T. D. 1991. *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life.* New York, NY: Free Press.

Greenberg, J. 1987. Reactions to procedural injustice in payment distributions: Do the means justify the ends? *Journal of Applied Psychology* 72 (1): 55–71.

Hayes, R. M., and S. Schaefer. 2000. Implicit contracts and the explanatory power of top executive compensation for future performance. *RAND Journal of Economics* 31 (2): 273–293.

Holmstrom, B. 1979. Moral hazard and observability. *The Bell Journal of Economics* 10 (1): 74–91.

Huber, V. L., M. A. Neale, and G. B. Northcraft. 1987. Judgment by heuristics: Effects of rate and rater characteristics and performance standards on performance-related judgments. *Organizational Behavior and Human Decision Processes* 40: 149–169.

Ittner, C. D., D. F. Larcker, and M. W. Meyer. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725–758.

Jacobs, R., and S. W. J. Kozlowski. 1985. A closer look at halo error in performance ratings. *The Academy of Management Journal* 28 (1): 201–212.

Jasso, G. 1978. On the justice of earnings: A new specification of the justice evaluation function. *The American Journal of Sociology* 83 (6): 1398–1419.

Klayman, J., and Y. W. Ha. 1987. Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review* 94 (2): 211–228.

Kravitz, D. A., and W. K. Balzer. 1992. Context effects in performance appraisal: A methodological critique and empirical study. *Journal of Applied Psychology* 77: 24–31.

Lipe, M. G., and S. E. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75 (3): 283–298.

Longenecker, C. O., H. P. Sims, and D. A. Gioia. 1987. Behind the mask: The politic of employee appraisal. *The Academy of Management Executive* 1 (3): 183–193.

Milgrom, P., and J. Roberts. 1992. *Economics, Organization and Management.* Englewood Cliffs, NJ: Prentice Hall.

Murphy, K. J., W. K. Balzer, M. C. Lockhart, and E. J. Eisenman. 1985. Effects of previous performance on evaluations of present performance. *Journal of Applied Psychology* 70 (1): 72–84.

Murphy, K. J., and P. Oyer. 2003. Discretion in executive incentive contracts: Theory and evidence. Working paper, University of Southern California and Stanford University.

Murphy, K. R., and J. N. Cleveland. 1995. *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives.* Thousand Oaks, CA: Sage.

Mussweiler, T., and F. Strack. 1999. Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology* 35 (2): 136–164.

Newcomb, T. 1931. An experiment designed to test the validity of a rating technique. *The Journal of Educational Psychology* 22: 279–288.

Nisbett, R. E., and T. D. Wilson. 1977. The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology* 35 (4): 250–256.

Nisbett, R. E., H. Zukier, and R. Lemley. 1981. The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology* 13: 248–277.

Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37 (1): 7–63.

Prendergast, C., and R. H. Topel. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37 (2–3): 355–365.

Prendergast, C., and R. H. Topel. 1996. Favoritism in organizations. *Journal of Political Economy* 104 (5): 958–978.

Tan, H. T., and K. Jamal. 2001. Do auditors objectively evaluate their subordinates' work? *The Accounting Review* 76 (1): 99–110.

Tan, H. T., and M. G. Lipe. 1997. Outcome effects: The impact of decision process and outcome controllability. *Journal of Behavioral Decision Making* 10: 315–325.

Wilks, T. J. 2002. Predecisional distortion of evidence as a consequence of real-time audit review. *The Accounting Review* 77 (1): 51–71.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1231–1253

American Accounting Association
DOI: 10.2308/accr-10039

Risk-Based Auditing, Strategic Prompts, and Auditor Sensitivity to the Strategic Risk of Fraud

Kendall Bowlin
The University of Mississippi

ABSTRACT: Under risk-based auditing, more (fewer) audit resources are allocated to accounts that are more (less) likely to be misstated. However, if auditors do not anticipate the strategic risk that arises when client managers anticipate auditors' risk-based resource allocations, undetected misstatements among ostensibly low-risk accounts could be more common than traditional risk assessment procedures suggest. Using a laboratory experiment, I find that participants assuming the auditor role in a stylized audit game do not naturally attune to strategic risks, but instead focus resources toward accounts with high non-strategic risk and away from ostensibly low-risk accounts. Manager-participants exploit these allocations by overriding the low-risk accounts more often than accounts with high non-strategic risk. However, auditor-participants who are asked to predict managers' expectations of, and responses to, audit resource allocations, devote additional resources to the low-risk accounts. These results are robust to the level of available audit resources.

Keywords: *audit resource allocation; strategic reasoning; fraud risk; risk-based auditing; experimental economics.*

Data Availability: *Data used in this study are available upon request.*

I. INTRODUCTION

The basic premise of risk-based auditing is that auditors should devote more resources to accounts that are likely to be misstated and fewer resources to those that are less likely to be misstated (Bell et al. 2005; Rittenberg and Schwieger 2005; Knechel 2007). This approach is expected to lead to more effective and efficient audits (e.g., Bell et al. 2005; Public Company

---

This paper is based on my dissertation at The University of Texas at Austin. I thank my dissertation committee members, Jeffrey Hales, Lisa Koonce, Paul Newman, Dale Stahl, and especially Steve Kachelmeier (chair) for their excellent guidance. I also acknowledge helpful comments from two anonymous reviewers, Billy Brewster, Brian Bratten, Margaret Christ, Michael Crawley, Jeremy Griffin, Jessen Hobson, Kathryn Kadous, Uday Murthy, Ed O'Donnell, Mark Peecher, Bernd Reichert, Ted Rodgers, Casey Schwab, Shankar Venkataraman, Jennifer Winchel, Mark Zimbelman, and workshop participants at The University of Mississippi, The University of Texas at Austin, the 2009 AAA Auditing Section Midyear Meeting, the 2009 AAA Annual Meeting, the 2009 Mid-South Accounting Research Consortium, and the 19th Symposium on Auditing Research at the University of Illinois at Urbana–Champaign. I am also grateful for financial support from the Eugene and Dora Bonham Endowment Fund, the McCombs School of Business Department of Accounting, and the Patterson School of Accountancy.

Editor's note: Accepted by Kathryn Kadous.

Submitted: July 2009
Accepted: January 2011
Published Online: April 2011

Accounting Oversight Board [PCAOB] 2007). However, if auditors do not accurately assess misstatement risk at the account level, audit resources will be misallocated, resulting in undetected misstatements (e.g., Kinney 2005; O'Donnell and Schultz 2005).

Auditors could wrongly assess misstatement risk by focusing on observable non-strategic risk factors that indicate certain accounts are more likely than others to be misstated and by failing to appreciate the attendant implications for unobservable strategic risks that arise when financial reporting managers anticipate that auditors will allocate resources based on those non-strategic risk factors (Fellingham and Newman 1985).[1] By fixating on non-strategic risk factors and by allocating resources accordingly, auditors could actually create opportunities for fraud among the ostensibly low-risk accounts.

The HealthSouth audit failure provides an example of such an outcome. Having apparently assessed fixed assets as a low-risk account, HealthSouth's auditors reportedly devoted relatively few resources to testing fixed assets, particularly to vouching smaller fixed asset additions. HealthSouth's management anticipated this strategy and exploited it by capitalizing certain expenses into the fixed asset accounts (Weil 2004).

The purpose of this study is to demonstrate this potential threat to the efficacy of risk-based auditing by examining the way in which auditors allocate resources among accounts that differ in non-strategic risk and by examining the strategic implications of these allocations for account-specific misstatement risk. I further examine whether prompting strategic reasoning reduces this threat. Although audit standards and extant research acknowledge that auditees might anticipate and exploit auditors' risk-based strategies (e.g., American Institute of Certified Public Accountants [AICPA] 2003; Peecher et al. 2007; Hoffman and Zimbelman 2009), few prior studies have proposed or tested interventions designed to enhance auditors' sensitivity to that possibility. An important exception is Hoffman and Zimbelman (2009), who find that prompting auditors to consider likely client responses to audit strategies leads to more effective modifications of a standard audit program in a single-account setting, where the risk of fraud specific to that account is known to be high.

My study builds on Hoffman and Zimbelman (2009) in two important ways. First, I explore the way in which auditors allocate limited pools of resources among client accounts that differ in non-strategic risk and whether client managers exploit those allocations. I also examine whether a strategic reasoning prompt enhances auditor sensitivity to the managers' potential exploitation of auditors' risk-based resource allocations and whether such prompts could thereby improve the effectiveness and efficiency of audit resource allocations. By comparison, Hoffman and Zimbelman (2009) focus on the use of prompts in improving auditor anticipation of and responses to specific modes of fraud when an account is already known to be high in fraud risk.

I address these questions with a laboratory experiment in which student participants assume the role of auditor or financial reporting manager and interact in a two-account audit setting, where the accounts differ in that one has higher non-strategic risk of misstatement than the other (hereafter the "high-risk" and "low-risk" accounts).[2] The manager chooses whether to override each account, and the auditor allocates a limited pool of resources between the two accounts. Although these accounts differ greatly in non-strategic risk, the auditor's Nash equilibrium resource allocations are much less extreme than the differential non-strategic risks would imply. That is, because of the low-risk account's lower non-strategic risk, the strategic manager expects that account to draw relatively little auditor attention. Therefore, the manager views it as a ripe opportunity for intentional misstatement, creating differential strategic risks that an auditor should consider along with the exogenous non-strategic risks. This between-account difference in non-strategic risk provides a within-subjects variable as part of the experimental design.[3]

---

[1] For example, the quality of the accounting system and internal controls, frequency and complexity of transactions and prevalence of management judgments are observable factors that could suggest that misstatements are more likely to occur within some accounts than in others.

[2] Experimental materials do not use terms such as auditor, manager, audit, fraud, etc., but instead use neutral terminology so that participants' choices were not influenced by their beliefs about how auditors and managers ought to behave (Haynes and Kachelmeier 1998). However, for clarity, I generally use the more meaningful terminology in this manuscript. For further expository convenience, I generally refer to participants simply as "auditors" and "managers," although all participants are students who assume those roles for the purposes of this study.

I also manipulate, between-subjects, whether auditors receive a strategic prompt. Auditors in the no-prompt condition are provided with the non-strategic risk probabilities. Auditors in the prompt condition are likewise informed of the non-strategic risk probabilities and are also prompted to consider the strategic implications of that information.[4] Specifically, each auditor in the prompt condition is asked to report his/her beliefs about how the manager expects the auditor to allocate resources and how the manager will likely respond to his/her expectations.

Prior research has demonstrated that resource constraints can affect the way in which auditors perceive and respond to risk (Houston 1999; Asare et al. 2000), which suggests that the constraints on audit resources could also affect the usefulness of strategic prompts. Therefore, I additionally manipulate the amount of resources available to the auditor in order to determine whether the effectiveness of the strategic prompt is robust to the degree of resource constraint. In the adequate resources condition, resources are limited to an amount that slightly exceeds that required by the Nash equilibrium, whereas in the excess resources condition, the level of available resources is well above the equilibrium level.

I find that participants in the auditor role tend to allocate fewer resources to the low-risk account than the high-risk account and participants in the manager role respond by overriding the low-risk account more often than the high-risk account. This results in more undetected misstatements in the ostensibly low-risk account than in the high-risk account, despite the high-risk account's higher level of non-strategic risk. However, compared to auditors in the no-prompt condition, prompted auditors tend to utilize more of their available resources, allocating significantly more resources to the ostensibly low-risk account but no more resources to the high-risk account. Thus, the strategic prompt does not simply result in a general increase in resource use, which could be inefficient, but instead enhances strategic reasoning, with the result that resources are deployed more effectively while maintaining efficiency. I further find that these results are robust to the level of resources available to the auditor, though auditors do utilize more resources overall within the excess resources condition.

These results demonstrate a potential threat to the effectiveness of risk-based auditing that has been largely overlooked in the accounting literature; that is, by failing to appreciate the strategic nature of fraud, auditors are unlikely to appropriately deploy their resources, resulting in failure to detect the intentional misstatements among ostensibly low-risk accounts that occur when managers anticipate that auditors will not focus resources on those accounts. However, the results further suggest that prompting auditors to form beliefs about managers' expectations of and responses to auditors' strategies enhances auditor sensitivity to strategic fraud risk, resulting in more effective deployment of resources. Therefore, standard-setters and audit firms could potentially improve audit effectiveness by incorporating such strategic prompts into auditing standards and other professional guidance. These results could be particularly important to the Public Company

---

[3] Phillips (1999) similarly examines auditor sensitivity to the presence of fraud in a multi-account setting, where misstatement risk differs between accounts. However, Phillips (1999) focuses on the eventual evaluation of audit evidence, whereas the current study is concerned with sensitivity to the strategic implications of differences in non-strategic risk during the risk assessment and planning phases of the audit.

[4] In practice, auditors likely learn of non-strategic risk probabilities through experience, professional or firm guidance, and analysis of client accounting systems.

Accounting Oversight Board (the "PCAOB"), which encourages a risk-based approach to integrated audits of internal controls and financial statements (PCAOB 2007).

In Section II, I describe the setting for this study as well as the game-theoretic predictions. I develop hypotheses and describe the experimental design and method in Sections III and IV, respectively. Section V presents analysis of the results, followed by concluding comments in Section VI.

II. SETTING

Game theoretic models of auditing (e.g., Shibano 1990; Bloomfield 1995; Newman et al. 2005) demonstrate that the risk of fraud is endogenous, in that the financial reporting manager's optimal reporting strategy depends on his/her expectations of the auditor's strategy. This study examines a multi-account audit setting constructed following the game-theoretic reasoning described in that prior analytic literature and adapted in prior behavioral research (e.g., Bloomfield 1997; King 2002; Fischbacher and Stefani 2007; Bowlin et al. 2009).

In this setting, the auditor is responsible for detecting misstatements within two financial statement accounts, $i \in$ (*low-risk*, *high-risk*), either of which could be misstated due to a non-strategic accounting system error or due to the financial reporting manager's strategic override of the accounting system. To that end, the auditor allocates a fixed pool of audit resources among the two accounts and his/her personal consumption.

Following Bloomfield (1997), the accounting system materially misstates each account with probability r_i, which represents each account's non-strategic risk. Although the value of r_i is common knowledge by assumption, neither the manager nor the auditor knows whether the accounting system will actually generate a misstatement prior to making their respective choices. Thus, simultaneous with the accounting system's outcome, the manager chooses whether to accept that outcome or to override the accounting system. If the manager chooses to override an account, that account will be misstated with certainty, but will not also be unintentionally misstated. Although multiple misstatements could occur within a single account in practice, my assumption here does not alter the strategic tension between the auditor and the manager but does allow a more straightforward analysis of the game and implementation in the laboratory.

Without knowledge of the manager's decisions, the auditor chooses the number of audit resources units, e_i, to devote to each account. The probability that the auditor will detect any existing misstatement in account i, p_i (the detection probability), is the product of e_i and an audit effectiveness multiplier, T, which represents the marginal increase in the auditor's probability of detecting a misstatement for each additional unit of resources devoted to a given account. Although this factor likely varies across accounts in practice, for the purposes of this study, T is set to 1 percent for all accounts. This means that, for each additional resource unit devoted to an account, the probability of detecting an existing misstatement increases by one percentage point. Accordingly, if the auditor allocates 100 units to account i (i.e., $e_i = 100$), the detection probability, p_i, is 100 percent. At the other extreme, if no resources are devoted to account i (i.e., $e_i = 0$), there is zero probability of detecting a misstatement in that account.

After the manager has chosen whether to override the two accounts and the auditor has allocated his/her resources, outcomes are determined randomly based on those choices and the non-strategic risk, r_i. Each player's payoff depends on these outcomes as described below.

Payoffs and the Nash Equilibrium

The auditor's total cost of resources is $\sum_i e_i K$, where K is the auditor's marginal utility of unused resources (i.e., cost per unit of resource). Consistent with King (2002), the total resources devoted to the two accounts is exogenously limited by an audit fee, F. Any resources that are not allocated to one of the accounts are consumed by the auditor. In addition to the utility garnered from unused resources, the auditor receives utility of Y for each account for completing the audit.[5] When undetected misstatements occur, the auditor also incurs a penalty, Z, which does not depend on whether the undetected misstatement was intentional or unintentional.

For each account, if the manager accepts the balance generated by the accounting system, his/her payoff is M_L when a misstatement does not occur or when a misstatement does occur but is not detected. When a misstatement is detected, the manager's payoff is $M_L - V_L$. On the other hand, when the manager overrides the accounting system, he or she receives M_H when the auditor does not detect the resulting intentional misstatement and $M_H - V_H$ when that misstatement is detected. I assume that the financial reporting manager prefers to override the accounting system when misstatements will not be detected, but prefers to accept the system's output when misstatements will be detected. That is, $M_H > M_L$, but $M_H - V_H < M_L - V_L$.

Solving for the equilibrium gives $q_i{*} = \frac{K - Zr_iT}{T(Z - Zr_i)}$ and $e_i{*} = \frac{M_H - M_L}{T(V_H - V_Lr_i)}$, as derived in Appendix A. Consistent with the logic of a mixed-strategy equilibrium (Fudenberg and Tirole 1991), the auditor's (manager's) optimal strategy is a function of the manager's (auditor's) payoff structure and the account-specific non-strategic risk, r_i, but does not depend on his/her own payoffs. Additionally, the equilibrium is unaffected by the amount of available resources. However, I assume that the amount of resources available to the auditor is sufficient to provide an equilibrium level of resources to both accounts. If resources were constrained to lower levels, the auditor would always allocate below-equilibrium resources to at least one account, thereby resulting in corner solutions, where the manager always overrides at least one of the accounts. Assuming a sufficient level of available resources allows me to focus on the more interesting circumstances with interior solutions.

Table 1 summarizes the game parameters, which are intended to capture the relative preferences of real-world auditors and managers. Note that the two accounts in this game differ only in the non-strategic risk parameter, r_i, which is set at 20 percent and 80 percent for the low-risk account and the high-risk account, respectively. Accordingly, the equilibrium is approximately ($e_1{*}$, $e_2{*}$, $q_1{*}$, $q_2{*}$) = (48, 52, 0.75, 0) for both resource conditions.

Under these parameters, the combination of the non-strategic risk of error and the strategic risk of fraud results in an equilibrium misstatement risk of 80 percent for each account, regardless of the non-strategic risk.[6] Therefore, the auditor will allocate similar amounts of resources to the two accounts, even though an unintentional misstatement is substantially more likely to occur in the high-risk account than the low-risk account.

Although the equilibrium misstatement risk is the same for both accounts, the auditor does allocate slightly more resources to the high-risk account in equilibrium. This occurs because the manager is penalized when an unintentional misstatement is detected. Since the manager cannot control the quality of the accounting system in this setting, this penalty for unintentional misstatements actually reduces the disincentive effect of the penalty for detected override. Therefore, as the quality of the accounting system decreases (i.e., as the non-strategic risk increases), the manager's incentive to override would increase were it not for the auditor's response of allocating more resources to the high-risk account. Although allocations of auditor effort could be more sharply skewed towards the high-risk account in practice, the equilibrium in this setting strengthens the ability of an experiment to test the theoretical premise that endogenous management reporting can offset differences in non-strategic risk.

---

[5] Although the value of Y does not affect the equilibrium solution to the audit game (as derived in Appendix A), a sufficiently high value of Y serves the practical purpose of ensuring non-negative payoffs to experimental participants. Therefore, for consistency, I include it in the auditor's utility function here.

[6] In other words, the auditor is made indifferent among all possible resource allocations when the manager chooses an override strategy that results in misstatement risk of 80 percent for each account.

TABLE 1

Game Parameters

Panel A: Parameters

	Adequate Resources	Excess Resources
F	101	181
K	0.12	0.12
Y	15	15
Z	15	15
T	0.01	0.01
$r_{low\text{-}risk}$	20%	20%
$r_{high\text{-}risk}$	80%	80%
M_H	15	15
M_L	8	8
V_H	15	15
V_L	2	2

Panel B: Equilibrium Choice Probabilities

	Adequate Resources	Excess Resources
$e_{low\text{-}risk}{}^*$	47.95	47.95
$e_{high\text{-}risk}{}^*$	52.24	52.24
$q_{low\text{-}risk}{}^*$	75.00%	75.00%
$q_{high\text{-}risk}{}^*$	0.00%	0.00%

This model incorporates a number of assumptions that simplify its implementation in the laboratory. First, although this game is meant to represent a fraud setting, the manager does not issue an accounting report on which the auditor opines, but instead chooses a hidden action (whether to override the accounting system), which is more often associated with misappropriation of firm assets than with misreporting in the analytic literature (Newman et al. 1996). Second, consistent with Anderson and Young (1988), I assume that detection of a misstatement in one account does not make a misstatement in the other account more likely to be detected. Third, I assume that the auditor's penalty for undetected misstatements does not differ depending on whether or not the misstatement is intentional. While, in practice, those penalties could differ, that difference would not affect the equilibrium predictions of auditor behavior. Game theory predicts that increasing the auditor's penalty for undetected fraud, while holding constant the penalty for undetected errors, results in a lower likelihood of fraud. However, the manager would continue to override accounts with lower non-strategic risk more often than those with higher non-strategic risk except as auditor penalties approach infinity. Finally, this model treats the audit as a one-shot game, though the same auditor and manager often interact repeatedly in practice, allowing each party to learn about the other's behavior. A multiperiod Bayesian analysis of the game could provide further analytic insights and predictions about auditor and manager behavior. Although each of these assumptions has important implications for predicted behavior, the purpose of this model is not to form precise predictions about auditor and manager behavior *per se*, but rather is to capture the strategic nature of the auditor-manager relationship in a multi-account setting and to allow a clear examination of auditor resource allocation strategies, manager responses, and the potential benefits of a strategic prompt.

III. HYPOTHESES

Effect of Non-Strategic Risk

The equilibrium solution to the multi-account audit game described in Section II is derived under the assumption that both the auditor and the manager are each able to anticipate the other's strategy and respond optimally to it. However, people often exhibit low levels of such strategic reasoning (e.g., Ochs 1995; Bloomfield 1997; Zimbelman and Waller 1999; Wilks and Zimbelman 2004; Fischbacher and Stefani 2007; Bowlin et al. 2009), and instead adopt intuitive decision rules when faced with cognitively challenging problems (Tversky and Kahneman 1986).[7] Adoption of intuitive decision rules is likely in the multi-account audit game due to the cognitive complexity of calculating the game's mixed-strategy equilibrium. Further, since the only observable difference between the two accounts in this setting is the non-strategic risk of misstatement, that risk will likely provide a salient focal point (Sugden 1995) upon which the auditor will construct an intuitive allocation strategy, directing substantially fewer resources to the low-risk account than the high-risk account in a manner contrary to the game theoretic prediction.

H1: Auditors will allocate a smaller percentage of resources to the low-risk account than to the high-risk account.

Manager Behavior

Allocating significantly more resources to the high-risk account, as predicted by H1, would be reasonable if managers do not anticipate or respond to that strategy, but such an allocation is ineffective if managers behave strategically. The game theoretic model in Section II predicts that the manager will anticipate the audit strategy predicted in H1 and will respond by overriding the low-risk account with a greater probability (75 percent) than the high-risk account (0 percent). While managers probably initially exercise limited strategic reasoning, they are likely to adopt a testing strategy (Bowlin et al. 2009) where they experiment by overriding each account. Therefore, their feedback will likely be informative about auditor behavior and indicate that overriding the high-risk account is detected by the auditor more often than overriding the low-risk account, leading managers to behavior that is consistent with the formal game theoretic predictions.

H2: Managers will override the low-risk account more often than the high-risk account.

Strategic Reasoning and the Effect of Strategic Prompts

Prior economics and accounting research has modeled strategic reasoning as a hierarchy of beliefs (e.g., Stahl and Wilson 1995; Zimbelman and Waller 1999; Camerer et al. 2004; Hoffman and Zimbelman 2009) in which strategic reasoning is categorized in levels according to the players' conjectures about their opponents' beliefs and likely behavior. Following Stahl and Wilson's (1995) model, Level-0 players make no assumptions about other players and haphazardly choose from available strategies. Players with Level-1 beliefs assume that all other players are Level-0 types. Since these Level-1 players assume their opponents make choices haphazardly, Level-1 players in mixed strategy games, such as my audit game, are likely to be responsive to components of their own utility functions. The behavior predicted in H1 is consistent with Level-1 beliefs. Level-2 players believe that all other players are a mixture of Level-0 and Level-1 players (Stahl and Wilson 1995; Camerer et al. 2004). Therefore, these players are more likely to be responsive to their opponents' utility functions, making them more strategically sophisticated.

---

[7] Wilks and Zimbelman (2004) define the auditor's strategic reasoning as the ability to anticipate the financial reporting manager's response to the auditor's strategy.

While prior research demonstrates that people typically attain low levels of strategic reasoning (Ochs 1995; Zimbelman and Waller 1999; Camerer et al. 2004; Fischbacher and Stefani 2007; Bowlin et al. 2009), few studies have tested interventions intended to raise the level of strategic reasoning. Hoffman and Zimbelman (2009) provide an important exception by examining the use of strategic prompts within an auditing context. In their study, professional auditors are asked to suggest modifications to a standard audit program after reviewing case information based on an actual fraud event. Before making these recommendations, participants in the strategic condition are required to explain how management might commit fraud in accounts receivable and how that fraud might be concealed if management anticipates the standard program. The authors find that auditors receiving this prompt recommend procedures that are more consistent with those suggested by a panel of expert auditors.

My study builds on Hoffman and Zimbelman (2009) in two important ways regarding the use of strategic prompts. First, their study is not interactive, meaning that the managers, who could anticipate audit strategies, are hypothetical. Therefore, their study cannot explore whether managers do anticipate and respond to the auditors' basic audit programs or whether they might further anticipate and respond to the auditors' modifications. An interactive setting allows one to demonstrate the need for enhanced strategic reasoning via examination of the strategic outcomes of the audit game. Second, Hoffman and Zimbelman (2009) consider a single-account setting where the risk of fraud is already known to be high. My study, on the other hand, takes a step back and examines the efficacy of a strategic prompt in improving the assessments of and responses to misstatement risk, which Hoffman and Zimbelman (2009) take as given.

A few prior studies in economics do examine the effects of strategic prompts in interactive settings with mixed results. For example, Croson (2000) finds that players are more likely to choose the dominant equilibrium strategies of free-riding in a public goods game and defection in a prisoners' dilemma game when players are required to make guesses about their opponents' strategies. However, Costa-Gomes and Weizsäcker (2008) find that participant choices in a series of one-shot 3 × 3 normal-form games are not affected by requiring participants to state beliefs about their opponents' choices before choosing their own strategies. Similarly, Ortmann et al. (2000) find that sender investments in a standard trust game do not differ when senders are asked to predict how much of firm product will be returned to them by the receivers.[8]

Differences in strategic complexity potentially explain these mixed results. Specifically, in public goods and prisoner dilemma games, where each player's optimal strategy is to defect regardless of the opponent's choice, a strategic prompt will likely highlight that dominant strategy, making its play more likely. However, the prompt is likely to be less effective in more strategically complex games like the trust game in Ortmann et al. (2000) and the normal-form games in Costa-Gomes and Weizsäcker (2008), where each player's optimal strategy depends on the opponent's choice, because predicting the opponent's choice will likely be difficult even in the presence of a strategic prompt.

Though an audit is strategically complex (Bloomfield 1995), a strategic prompt is likely to be more effective in my multi-account setting because judgments about risk can be made by contrasting the two accounts rather than by evaluating each one independently (Bhattacharjee et al. 2007). Even if an auditor receives a strategic prompt, it is still difficult to assess the absolute probability that the manager will override any particular account. However, the auditor could more easily predict that the manager is more likely to override the low-risk account than the high-risk account. That is, the same comparison of non-strategic risks that provides a focal point for the intuitive allocation strategy predicted in H1 will also help prompted auditors recognize that client managers are likely to believe auditors will focus fewer resources on low-risk accounts and are more likely to misreport those accounts. Therefore, auditors in the prompt condition are likely to devote more resources to low-risk accounts than unprompted auditors.

---

[8] In a trust game, a sender receives an endowment and chooses the portion of the endowment to invest with a receiver. The amount invested is multiplied and the receiver then chooses a portion of the product to return to the sender. Game theory predicts that the receiver would choose to maximize his/her profit by keeping the entire product and that the sender will, therefore, not invest any of his/her endowment. However, contrary to these economic predictions, prior experimental studies find that senders often invest at least part of the endowment and that receivers often return some of the product to the senders (e.g., Berg et al. 1995).

> **H3a:** Auditors receiving the strategic prompt will allocate more resources to the low-risk account than auditors not receiving such a prompt.

If the prompt does indeed enhance auditors' sensitivity to the strategic risk of fraud, auditors will recognize that managers are more likely to override the low-risk account than the high-risk account. Therefore, while the prompt is expected to result in additional resources being devoted to the low-risk account, the prompt will not likely result in the allocation of more resources to the high-risk account.

> **H3b:** Auditors receiving the strategic prompt will not allocate more resources to the high-risk account than auditors not receiving such a prompt.

Effect of Resource Availability

Prior research suggests that the way in which auditors perceive and respond to risk could depend on the degree to which audit resources are constrained (Glover 1997; Houston 1999). Tight resource constraints are likely to induce constraint stress, which occurs when one feels externally prevented from doing what is desired or necessary (Schuler 1980). Prior auditing research indicates that when auditors face constraint stress, they often respond by mentally filtering away information that seems less relevant to the judgment at hand (e.g., Glover 1997; Asare et al. 2000).[9]

When auditors in strategic settings face tight resource constraints, they could similarly respond by filtering away information that seems less relevant to the required judgments. However, in the setting described in Section II, the information that seems less relevant would likely be the information about the managers' incentives and potential strategic behavior, the use of which is likely to seem more ambiguous and less obvious than the information about non-strategic risk, which clearly differs between accounts. Contrary to auditors in the no-prompt condition, who are less likely to comprehend the strategic implications of the game, auditors in the prompt condition are likely to recognize the strategic nature of the game, but could be tempted to filter that information when making the allocation decision if resources are more tightly constrained. Auditors with excess resources are less likely to suffer from constraint stress and, therefore, would be less likely to filter strategic information. Consequently, the strategic prompt could be more effective under the excess resources condition than in the adequate resources condition.

On the other hand, the opposite prediction is not unreasonable. For example, economics research has shown that introducing excess cash into a market can result in large price bubbles even when fundamental value is well-known (Caginalp et al. 2001). Similarly, when auditors have access to excess resources, they could avoid cognitively challenging allocation problems by devoting large amounts of resources to both accounts (relative to the equilibrium) and still retain some resources as profit. Under these circumstances, auditors are not likely to attune as closely to the differences in risk between the two accounts. If this is true, then having greater resources could make a prompt less effective.

---

[9] Glover (1997) finds that auditors filter less relevant information when they face time limits during the conduct of substantive audit procedures. Asare et al. (2000) extend Glover (1997) by operationalizing time pressure with constrained budgets rather than time limits. They find evidence suggesting that, during the planning phase, auditors similarly filter information when choosing audit procedures under tighter time budget constraints. My resource manipulation is analogous to Asare et al.'s (2000) budget manipulation.

These two lines of reasoning support predictions of opposite ordinal interactions between the prompt and resource variables, suggesting that the effect of the strategic prompt would not be robust across differing levels of available audit resources. Given that it is unclear which of these two possible effects will dominate or whether both effects may operate simultaneously within a population, I state the fourth hypothesis in the null form:

> **H4:** The effect of the strategic prompt on the amount of resources allocated to the low-risk account will not differ between the adequate and excess resources conditions.

IV. EXPERIMENTAL METHOD

Environment and Procedures

The experiment was conducted in a controlled laboratory environment with 132 volunteer upper-division accounting students as participants. Prior to beginning the experiment, participants were provided with written instructions that the experimenter read aloud. The instructions explained that each player would be anonymously assigned to one of two player types, a "Chooser" (analog to a financial reporting manager) or a "Guesser" (analog to an auditor), with each Chooser paired with one Guesser for 20 rounds of the audit game described in Section II.[10] The instructions familiarized participants with the setting of the game, the choices to be made by each player type, and their respective payoffs for each of the possible game outcomes.

To enhance comprehension, the instructions characterized the game as a production process with two machines (analog to the accounting system) that each deposit 100 marbles into its output bin. Of the 100 marbles placed in each bin, one will be an odd color (analog to a misstatement) if the Chooser overrides the machine or if the machine places an odd-colored marble in its bin by chance. The instructions explained that the Guesser may search for odd-colored marbles by specifying the number of marbles to be drawn from each bin up to a maximum total of 101 and 181 marbles in the adequate and excess resource conditions, respectively.

The game was implemented using z-Tree software (Fischbacher 2007) on networked laboratory computers. After the experimenter finished reading the instructions, participants received supplemental instructions via z-Tree, which summarized the paper-based instructions and also described the strategic prompt to auditors in the prompt condition. After reading these supplemental instructions, participants took a true-false quiz to help ensure their understanding of the instructions.[11]

After all participants completed the quiz, the computer began the first round of the game described above. In all conditions, the managers chose whether to override each of the two accounts, and the auditors chose how much of their resources would be allocated to each account and how much would remain unallocated. However, before making their allocation decisions, auditors in the prompt condition responded to two belief-elicitation questions that made up the strategic prompt described in more detail below. When all decisions had been entered, the computer randomly generated outcomes conditional on participant choices and the account-specific non-strategic risk. The computer then awarded points according to the parameters in Table 1. After receiving feedback reports, the game was repeated as described above for another 19 rounds.

---

[10] Although the player labels were intended to reduce potential demand effects, the "Guesser" label could potentially suggest that participants in that role should allocate resources at random. However, results reported in Section V suggest that this is not the case. Additionally, in their responses to open-ended post-experimental questions about their decision-making, nearly all participants describe allocation strategies that are more sophisticated than a simple guessing strategy.

[11] To help ensure that the instructions were understood, z-Tree provided an explanation of the relevant portion of the instructions when a participant answered a question incorrectly. The participant then answered that question again before moving on to the next question.

Feedback and Compensation

After the computer determined the outcomes at the end of each round, each participant received feedback for each account. The auditor's feedback only reported his/her own allocation choice and whether the audit had detected a misstatement. If no misstatement was detected, the auditor was not informed about whether a misstatement had occurred but had been undetected. Similarly, the manager's feedback reported his/her own choices and whether a misstatement had been detected but did not report the amount of audit resources devoted to each account. The incomplete nature of this feedback reflects an institutional reality that auditors cannot know with certainty that an account balance is correctly stated unless that account is fully audited (Anderson and Young 1988), and that managers cannot know with certainty the auditor's overall level of effort or how the auditor allocates his/her resources (Bowlin et al. 2009).

Note that the auditor only learned whether s/he detected a misstatement and not whether a detected misstatement occurred by chance or due to manager override. This is consistent with Statement on Auditing Standards No. 99 (SAS 99), which suggests that auditors sometimes have difficulty determining whether a detected misstatement is intentional or unintentional (AICPA 2002, footnote 4).

Consistent with the tenets of experimental economics, participants were privately compensated at the end of each session. Participants received a show-up fee of $5.00 plus 6 cents per point earned. The average participant earned approximately $28 for sessions lasting no more than 90 minutes.

Manipulation

In order for an auditor to choose an effective and efficient allocation strategy, his/her understanding of the audit should include the possibility that the manager will anticipate and attempt to exploit the resource allocation predicted in H1. However, a prompt that simply asks the auditor to predict the manager's strategy may not sufficiently enhance the auditor's understanding of managerial incentives. That is, to the extent that auditors naturally engage in Level-1 reasoning, a prompt that merely elicits the auditor's beliefs about the manager's likely choice might only elicit Level-1 priors (that the manager will behave randomly).

Therefore, I test a stronger strategic prompt that begins by asking an auditor to predict the manager's beliefs about the auditor's allocation strategy and then further asks the auditor to predict the manager's response to those beliefs. Eliciting both types of beliefs is more likely to help auditors develop a strategically sophisticated understanding of the audit environment.

Accordingly, at the beginning of each round, auditors in the prompt condition were required to state their beliefs about the managers' expectations of and likely responses to auditors' resource allocations prior to specifying the amount of resources to be allocated to each account. Specifically, and continuing the marble and bin analogy, each auditor in the prompt condition entered his/her beliefs about the number of marbles the manager expected the auditor to draw from each bin. Then, using a Likert scale, the auditor entered his/her beliefs about the likelihood that the manager would override each machine. To maintain *ceteris paribus* conditions, the managers were not made aware of this prompt and auditor compensation was not dependent on their responses to it.[12]

The level of available resources was also manipulated between subjects. Continuing the marble and bin analogy, auditors in the adequate resources condition could draw up to 101 of the 200 total marbles in the two bins, which is sufficient to play the equilibrium strategy described in Section II. However, auditors in the excess resources condition could draw as many as 181 total marbles from the two bins. The level of resources available to the auditor was common information to both auditors and managers.

V. RESULTS

To construct independent observations, each participant's choices are averaged across all 20 rounds, such that each participant provides a single observation of auditor or manager behavior. For each account, each auditor observation represents the average number of resource units, e_i, allocated to that account. Each manager observation for each of the two accounts represents the percentage frequency with which the manager chose to override that account. Panels A, B, and C of Table 2 summarize the mean number of resource units that auditors utilize in total and specifically devote to the low-risk and high-risk accounts, while Panels A and B of Table 4 summarize manager behavior under each experimental condition.[13]

Preliminary Analysis of Auditor Behavior

Pooling across resource conditions, auditors utilize a mean of 110.0 resource units when they receive the prompt, compared to an average of 98.6 units when they are not prompted. In addition, auditors utilize 122.5 resource units in the excess resources condition, but only utilize an average of 84.7 units in the adequate resources condition. The significant main effects of *PROMPT* ($t_{62} = 1.87$, one-sided $p = 0.033$) and *RESOURCES* ($p < 0.001$) in the three-factor, repeated-measures ANOVA on audit resource allocations in Panel A of Table 3 confirm that both prompting auditors to consider the strategic nature of the audit and making more resources available to the auditor result in the utilization of significantly more audit resources overall.

Effect of Non-Strategic Risk (H1)

Recall that H1 predicts that auditors will allocate more resources to the high-risk account relative to the low-risk account. Consistent with H1, results tabulated in Panels B and C of Table 2 show that more resources are allocated to the high-risk account than the low-risk account under all treatments. Pooling across prompt conditions, in the excess resources condition, auditors allocate an average of 68.8 and 53.7 resource units to the high- and low-risk accounts, respectively. Similarly, auditors allocate 55.4 units to the high-risk account and 29.2 to the low-risk account in the adequate resources condition. A significant main effect of *ACCOUNT* ($p < 0.001$, Table 3, Panel A)

---

[12] Not compensating auditors based on their responses to the prompt has the added benefit of maintaining the incentive dominance of the auditors' resource allocation decisions. However, since participants were not compensated based on the quality or accuracy of their responses to the strategic prompt, those responses may not reliably report their true beliefs while nevertheless enhancing strategic understanding of the audit setting. In fact, analyses of those responses discern few relationships among them aside from a positive correlation between auditors' beliefs about the number of resource units managers expect auditors to allocate to each and the amount of resources auditors actually allocate to each account.

[13] Levene tests for equality of variance (Levene 1957) suggest that there is significant heterogeneity among the cell variances reported in Panels A and B of Table 2. This heterogeneity can be reduced by converting the data to ranks (Kachelmeier and Messier 1990). However, inferences based on the analysis of rank-transformed data do not qualitatively differ from inferences based on the analysis of raw data. Therefore, for consistency across analyses, the results discussed in the following sections are based on the raw data.

TABLE 2

Auditor Resource Allocations by Treatment

Panel A: Mean Total Audit Resources Utilized

	No-Prompt	Prompt	Combined
Excess Resources	114.8	130.8	122.5
	(42.6)	(32.6)	(38.4)
	n = 18	n = 17	n = 35
Adequate Resources	78.0	90.1	84.7
	(24.9)	(8.9)	(18.7)
	n = 14	n = 17	n = 31
Combined	98.6	110.0	
	(39.9)	(31.3)	
	n = 32	n = 34	

Panel B: Mean Resources Allocated under the Adequate Resources Condition

	No-Prompt	Prompt	Combined
High-Risk Account	53.7	56.9	55.4
	(23.8)	(14.0)	(18.8)
	n = 14	n = 17	n = 31
Low-Risk Account	24.3	33.3	29.2
	(20.0)	(14.1)	(17.1)
	n = 14	n = 17	n = 31
Combined	39.0	45.1	
	(26.1)	(18.3)	
	n = 28	n = 34	

Panel C: Mean Resources Allocated under the Excess Resources Condition

	No-Prompt	Prompt	Combined
High-Risk Account	68.1	69.5	68.8
	(25.2)	(16.6)	(21.1)
	n = 18	n = 17	n = 35
Low-Risk Account	46.7	61.3	53.7
	(26.2)	(24.6)	(26.2)
	n = 18	n = 17	n = 35
Combined	57.4	65.4	
	(27.8)	(21.1)	
	n = 36	n = 34	

Standard deviation in parentheses.

supports H1, confirming that the overall difference in resources devoted to the high-risk and low-risk accounts is significant.

Manager Behavior (H2)

Consistent with H2 and as presented in Table 4, managers override the low-risk account more often than the high-risk account under both prompt conditions. Managers in the adequate

TABLE 3

Analysis of Audit Resource Allocations

Panel A: Three-Factor Repeated Measures ANOVA on Audit Resource Allocations

	df	SS	F	p-value
Between-subjects effects:				
PROMPT	1	1,611	3.49	0.066
RESOURCES	1	12,262	26.59	<0.001
PROMPT × *RESOURCES*	1	31	0.07	0.795
Between-subjects error	62	28,589		
Within-subjects effects:				
ACCOUNT	1	13,914	32.32	<0.001
ACCOUNT × *PROMPT*	1	746	1.73	0.193
ACCOUNT × *RESOURCES*	1	1,117	2.59	0.112
ACCOUNT × *PROMPT* × *RESOURCES*	1	118	0.27	0.603
Within-subjects error	62	26,695		

Panel B: Two-Factor Repeated Measures ANOVA on Resources Allocated to Low-Risk Account

	df	SS	F	p-value
PROMPT	1	2,275	4.77	0.033
RESOURCES	1	10,391	21.77	<0.001
PROMPT × *RESOURCES*	1	136	0.28	0.596
Error	62	42,010		

Panel C: Two-Factor Repeated Measures ANOVA on Resources Allocated to High-Risk Account

	df	SS	F	p-value
PROMPT	1	82	0.20	0.658
RESOURCES	1	2,988	7.21	0.009
PROMPT × *RESOURCES*	1	14	0.03	0.856
Error	62	25,690		

Variable Definitions:
Dependent Variable = auditor's mean allocation of resource units;
ACCOUNT = dichotomous factor to differentiate between the high-risk and low-risk accounts;
PROMPT = dichotomous factor to differentiate between treatments in which auditors receive the strategic prompt and those in which auditors do not receive the strategic prompt; and
RESOURCES = dichotomous factor to differentiate between adequate and excess levels of auditor resources.

resources condition override the low-risk account 67.4 percent of the time, while overriding the high-risk account 47.6 percent of the time. Similarly, in the excess resources condition, managers override the low-risk and high-risk accounts 49.9 percent and 35.1 percent of the time, respectively. A significant main effect of *ACCOUNT* ($t_{62} = 4.62$, one-tailed $p < 0.001$) in the three-factor, repeated-measures ANOVA reported in Table 5 indicates that managers override the low-risk account more often than the high-risk account, suggesting that managers are responsive to the auditors' strategies of allocating larger amounts of resources to the high-risk account than the low-risk account.

TABLE 4

Manager Override Frequency by Treatment

Panel A: Mean Manager Frequency of Override under the Adequate Resources Condition

	No-Prompt	**Prompt**	**Combined**
High-Risk Account	49.3%	46.2%	47.6%
	(29.3%)	(24.3%)	(26.2%)
	n = 14	n = 17	n = 31
Low-Risk Account	70.7%	64.7%	67.4%
	(22.4%)	(19.2%)	(20.6%)
	n = 14	n = 17	n = 31
Combined	60.0%	55.4%	
	(27.8%)	(23.5%)	
	n = 28	n = 34	

Panel B: Mean Manager Frequency of Override under the Excess Resources Condition

	No-Prompt	**Prompt**	**Combined**
High-Risk Account	35.0%	35.3%	35.1%
	(21.2%)	(16.4%)	(18.8%)
	n = 18	n = 17	n = 35
Low-Risk Account	51.9%	47.6%	49.9%
	(21.8%)	(22.6%)	(22.0%)
	n = 18	n = 17	n = 35
Combined	43.5%	41.5%	
	(22.9%)	(20.4%)	
	n = 36	n = 34	

Standard deviation in parentheses.

Untabulated results of a follow-up logistic regression of manager override decisions on *ACCOUNT*, *PROMPT*, *RESOURCES*, and an experience variable (*ROUND*), along with all possible interactions, indicates that the manager sensitivity to *ACCOUNT* is initially limited and develops with experience. Specifically, as participants in the manager role gain experience, they become increasingly more likely to override the low-risk account than the high-risk account (Wald $\chi^2 = 12.34$, $p < 0.001$).

Effect of Strategic Prompt (H3a and H3b)

Theory suggests that when auditors receive the strategic prompt, they will be more likely to recognize that managers will anticipate that fewer resources will be allocated to the low-risk account than to the high-risk account and that managers will likely view the low-risk account as a particularly attractive opportunity for fraud. Finding, as predicted in H3a and H3b, that the prompt increases audit resources devoted to the low-risk account without a similar increase in the allocation to the high-risk account would support the underlying argument that a strategic prompt enhances auditors' strategic understanding of the audit and of misstatement risk rather than simply prompting a general deployment of additional resources.

Observed resource allocations, presented in Table 2, are consistent with these predictions. Specifically, within the adequate (excess) resources condition, auditors allocate a mean of 33.3 and

TABLE 5

Three-Factor Repeated Measures ANOVA on Manager Override Frequency

	df	SS	F	p-value
Between-subjects effects:				
PROMPT	1	0.0352	0.67	0.417
RESOURCES	1	0.7605	14.40	<0.001
PROMPT × *RESOURCES*	1	0.0053	0.10	0.751
Between-subjects error	62	3.2736		
Within-subjects effects:				
ACCOUNT	1	0.9803	21.30	<0.001
ACCOUNT × *PROMPT*	1	0.0115	0.25	0.619
ACCOUNT × *RESOURCES*	1	0.0232	0.50	0.480
ACCOUNT × *PROMPT* × *RESOURCES*	1	0.0006	0.01	0.911
Within-subjects error	62	2.8536		

Variable Definitions:
Dependent Variable = manager's percentage frequency of choosing "override";
ACCOUNT = dichotomous factor to differentiate between the high-risk and low-risk accounts;
PROMPT = dichotomous factor to differentiate between treatments in which auditors receive the strategic prompt and those in which auditors do not receive the strategic prompt; and
RESOURCES = dichotomous factor to differentiate between adequate and excess levels of auditor resources.

24.3 (61.3 and 46.7) resource units to the low-risk account in the prompt and no-prompt conditions, respectively, while auditors allocate 56.9 and 53.7 (69.5 and 68.1) units to the high-risk account in the prompt and no-prompt conditions, respectively. While this pattern of means is consistent with H3a and H3b and suggests an ordinal interaction, the *ACCOUNT* × *PROMPT* effect in Table 3, Panel A, is only marginally significant ($t_{62} = 1.32$, one-tailed $p = 0.096$).[14] However, since ANOVA is not always powerful enough to detect significant ordinal interactions (Buckless and Ravenscroft 1990), I separately analyze the allocations of resources to the low-risk and high-risk accounts to provide further evidence that the additional resources used under the *PROMPT* condition are allocated in a manner consistent with an enhanced understanding of the strategic nature of misstatement risk. Panel B of Table 3 indicates that the increase in resources allocated to the low-risk account under the prompt condition is significant ($t_{62} = 2.18$, one-sided $p = 0.016$), which supports H3a. However, consistent with H3b, *PROMPT*'s effect on the amount of resources allocated to the high-risk account is not significant ($p = 0.658$, Table 3, Panel C). These results suggest that the prompt enhances auditor sensitivity to the strategic risk of fraud and consequently improves the effectiveness of auditor resource allocations without diminishing efficiency.[15]

Supplemental Analysis of Undetected Misstatements

The untabulated results of logistic regression of manager override behavior, described above, also indicates that the three-way interaction among *ACCOUNT*, *PROMPT*, and *ROUND* is negative and marginally significant (Wald $\chi^2 = 2.75$, $p = 0.098$), implying that the strategic prompt is associated with a growing downward pressure on managers' tendencies to override the low-risk account versus the high-risk account. Now, recall that auditors in the prompt condition allocate more resources to the low-risk account than unprompted auditors but no more resources to the high-risk account. Taken together, these findings suggest that, especially in later rounds, the prompt will lead to a larger decrease in the number of undetected misstatements in the low-risk account than in the high-risk account. The results reported in Tables 6 and 7 suggest this pattern of outcomes.

---

[14] McNeil et al. (1996, 137–139) discuss the rationale for one-tailed tests of directionally predicted interactions.

[15] An untabulated regression of the amount of resources allocated to the low-risk account during the last ten rounds of the game on *PROMPT* and *RESOURCES*, and the detection of manager override during the first ten rounds indicates a positive relationship between allocations to the low-risk account and previously detected override behavior. However, this learning does not appear to moderate the effects of *PROMPT* or *RESOURCES* on resources allocated to the low-risk account.

Table 6 summarizes, by treatment, the mean cumulative undetected misstatements that occur during the last ten rounds of the audit game. In both resource conditions, there are fewer undetected misstatements in the low-risk account in the prompt condition than in the no-prompt condition. Under adequate (excess) resources, there are a mean of 6.50 and 5.00 (3.78 and 2.12) cumulative undetected misstatements in the no-prompt and prompt conditions, respectively. However, there is little difference in the number of undetected misstatements in the high-risk account between the prompt and no-prompt conditions. Specifically, under the adequate (excess) resources condition, the cumulative undetected misstatements in the high-risk account average 4.07 and 3.94 (2.61 and 2.71) in the no-prompt and prompt conditions, respectively. A significant *ACCOUNT* × *PROMPT* interaction ($p = 0.048$, Table 7) confirms that the prompt has a larger effect on the number of undetected misstatements in the low-risk account than in the high-risk account.

TABLE 6

Cumulative Late-Round Undetected Misstatements by Treatment

Panel A: Mean Cumulative Late-Round Undetected Misstatements under Adequate Resources

	No-Prompt	Prompt	Combined
High-Risk Account	4.07	3.94	4.00
	(2.79)	(2.11)	(2.39)
	n = 14	n = 17	n = 35
Low-Risk Account	6.50	5.00	5.68
	(3.16)	(2.09)	(2.69)
	n = 14	n = 17	n = 31
Combined	5.29	4.47	
	(3.17)	(2.14)	
	n = 28	n = 34	

Panel B: Mean Cumulative Late-Round Undetected Misstatements under Excess Resources

	No-Prompt	Prompt	Combined
High-Risk Account	2.61	2.71	2.66
	(2.83)	(1.99)	(2.42)
	n = 18	n = 17	n = 35
Low-Risk Account	3.78	2.12	2.98
	(2.82)	(2.39)	(2.72)
	n = 18	n = 17	n = 35
Combined	3.19	2.41	
	(2.85)	(2.19)	
	n = 36	n = 34	

Standard deviation in parentheses.

TABLE 7

Three-Factor Repeated Measures ANOVA on Undetected Misstatements

	df	SS	F	p-value
Between-subjects effects:				
PROMPT	1	20.87	2.62	0.111
RESOURCES	1	140.81	17.65	<0.001
PROMPT × *RESOURCES*	1	0.01	0.00	0.974
Between-subjects error	62	494.56		
Within-subjects effects:				
ACCOUNT	1	33.79	6.90	0.011
ACCOUNT × *PROMPT*	1	19.96	4.08	0.048
ACCOUNT × *RESOURCES*	1	17.30	3.53	0.065
ACCOUNT × *PROMPT* × *RESOURCES*	1	0.30	0.06	0.804
Within-subjects error	62	303.49		

Variable Definitions:
Dependent Variable = session total number of undetected misstatements;
ACCOUNT = dichotomous factor to differentiate between the high-risk and low-risk accounts;
PROMPT = dichotomous factor to differentiate between treatments in which auditors receive the strategic prompt and those in which auditors do not receive the strategic prompt; and
RESOURCES = dichotomous factor to differentiate between adequate and excess levels of auditor resources.

While these results provide less direct evidence of the prompt's effect on auditor strategic reasoning than analysis of auditor resource allocations, examining the strategic prompt's effects on the level of undetected misstatements is nevertheless important, in that doing so provides broader insights into the prompt's direct and indirect effects on strategic behavior and audit outcomes. That is, the prompt's downward pressure on the managers' tendencies to override the low-risk account is likely due to the managers experiencing the auditors' prompt-induced attention to that account.

Also note in Panels A and B of Table 6 that there are more undetected misstatements in the low-risk account than the high-risk account in all treatments, except the prompt-excess resources treatment. The significant main effect of *ACCOUNT* (p = 0.011, Table 7) confirms that, in general, more undetected misstatements occur in the low-risk account than the high-risk account. This is an important observation for two reasons. First, it further demonstrates that an auditor's failure to understand the strategic nature of an audit could result in a misallocation of resources and more audit failures. Second, because managers do not fully exploit the auditors' below equilibrium allocations to the low-risk account by always overriding that account, as game theory would predict, it is important to show that the high-risk account is not, in fact, a riskier account, in which case allocating significantly more resources to that account than the low-risk account would be reasonable. Instead, these results indicate that allocating additional resources to the low-risk account is economically preferable in this setting.

Effect of Resource Availability (H4)

As discussed in Section III, prior research suggests that the level of resources available to an auditor could potentially moderate the strategic prompt's effect on resource allocations, though it is unclear whether increasing available resources would increase or decrease the effectiveness of the strategic prompt. However, an insignificant ANOVA interaction term (p = 0.596, Table 3, Panel B) indicates that the effect of *PROMPT* on the amount of resources allocated to the low-risk account is robust to the level of resources available to the auditor, providing no basis for rejecting H4.[16]

Although resource availability does not interact with the strategic prompt to affect the amount of resources allocated to the low-risk account, the main effect of *RESOURCES* ($p < 0.001$, Table 3, Panel A) indicates that the availability of resources does affect the overall use of them, which is inconsistent with the game theoretic predictions in Section II that the level of available audit resources will not affect auditor strategy. Pooling across prompt conditions, auditors allocate 53.7 resource units to the low-risk account in the excess resources condition, but in the adequate resources condition, auditors devote only 29.2 units to that account. Similarly, the amount of resources allocated to the high-risk account decreases from 68.8 units in the excess resources condition to 55.4 units in the adequate resources condition.

In total, a mean of 122.5 resource units are deployed in the excess resource condition, as compared to 84.7 units in the adequate resources condition. There are two potential explanations for this overall difference. First, the distribution of total deployed resources may have been drawn from equivalent distributions under both resource conditions, with the exception that the distribution is more truncated under the adequate resource condition where observations group at the adequate resources constraint of 101 resource units. Under this explanation, the amount of resources auditors desire to deploy would not be affected by resource availability and the frequency with which auditors deploy at least, say, 100 resource units would not differ between resource conditions. Alternatively, making more resources available to the auditor could result in a distributional shift, as auditors are able to experiment along a broader range of resources. In this case, one would find that auditors deploy at least 100 resource units more often when more resources are available.

Untabulated analysis supports the second explanation. The frequency with which auditors deploy at least 100 total resource units is greater in the excess resources condition than in the adequate resources condition (75.6 percent and 52.3 percent, respectively). An ANOVA of the frequency with which auditors use at least 100 resource units on *PROMPT* and *RESOURCES* indicates that this difference is significant ($F_{1,62} = 10.15$, $p = 0.002$), suggesting that constraining audit resources not only limits the amount of resources available for use but perhaps also constrains the amount of resources the auditors desire to use.

VI. CONCLUSION

Risk-based auditing is an intuitive approach that focuses more resources on accounts deemed to be riskier and fewer resources on accounts that seem less risky. However, if auditors focus on non-strategic risks and overlook the strategic risks that arise when managers anticipate risk-based allocations of audit resources, opportunities for fraud can arise among ostensibly low-risk accounts. The purpose of this study is to demonstrate this potential threat to the efficacy of risk-based auditing and to investigate whether prompting auditors to consider financial reporting managers' likely expectations of and responses to auditor strategies can enhance auditors' sensitivity to the strategic risk of fraud.

I conduct a laboratory experiment in which student participants assume the roles of auditors and financial reporting managers and interact in a multi-account audit setting, which reflects the essential strategic tension that exists between real-world auditors and managers. Results indicate that participants in the auditor role fail to appreciate the strategic nature of the multi-account audit game and, therefore, allocate fewer of their available resources to accounts with relatively low non-strategic risk. Participants in the manager role exploit this behavior by overriding this low-risk account more often than the high-risk account. These behaviors jointly result in more undetected misstatements within the low-risk account than the high-risk account. However, I also find that prompting auditor-participants to consider their opponents' expectations of and responses to audit resource allocations results in auditor-participants allocating more resources to the low-risk account without a similar change in resources allocated to the high-risk account. In later rounds of the game, these increased resources lead to fewer undetected misstatements in the low-risk account but no significant change in the number of undetected misstatements in the high-risk account. These results suggest that a strategic prompt can enhance auditor sensitivity to the strategic risk of fraud and thereby improve audit effectiveness without reducing efficiency. Finally, I also find that the relative effect of the strategic prompt is robust across levels of available resources, although increasing the amount of resources available to the auditor-participants results in more resources being used.

---

[16] Likewise, Table 4, Panel B indicates that the amount of resources allocated to the high-risk account is not significantly affected by the interaction of the strategic prompt and resource availability.

These results suggest implications for both audit firms and audit standard-setters, including the AICPA's Auditing Standards Board and the PCAOB. While current auditing standards generally promote risk assessment procedures that are intended to focus auditor attention toward observable factors indicating that certain financial statement accounts are more likely to be misstated than others (e.g., AICPA 2002, 2006), the results of this study suggest that audit procedures that prompt auditors to consider the possibility that financial reporting managers will anticipate and exploit auditor's risk assessments could result in fewer undetected misstatements. As an example of such a strategic prompt, audit standards or professional guidance could require that audit teams specifically discuss whether there are any client accounts that management would expect auditors to treat as low-risk. If the auditor determines that there are such accounts, standards could require the auditor to take the additional step of considering whether management is likely to attempt to exploit the expected lack of auditor attention. If auditors identify such accounts and determine that fraud attempts are sufficiently likely within them, they could then follow Hoffman and Zimbelman (2009) by considering the specific ways in which fraud could be perpetrated and concealed within those accounts and then designing procedures to overcome such concealment.

A limitation of this study is that the experimental design does not disentangle the marginal effects of eliciting auditors' beliefs about managers' expectations of auditor strategies and eliciting auditors' beliefs about managers' likely reporting choices. Though a broader purpose of this study is to examine whether strategic prompts, in general, can enhance auditors' strategic reasoning and sensitivity to fraud risk, disentangling these effects could be important to audit firms and standard-setters as they structure audit standards and guidance.

Consistent with the traditions of experimental economics (Kachelmeier and King 2002), this study attempts to capture the strategic essence of auditing using a relatively stark but tractable setting. As such, I abstract away several institutional features of real-world auditing that may affect auditor and manager choices, leaving them as potential avenues for future research. For example, the amount of resources available to a real-world auditor is likely to be endogenously determined through negotiation between the auditor and client. Additionally, in my setting, the manager is equally capable of intentionally misstating each financial statement account, though, in practice, certain low-risk accounts may not be as easily manipulated. Finally, I assume that auditor penalties do not differ between undetected fraud and errors, and while this does not impact the economic predictions of auditor behavior, such differences could behaviorally affect auditor sensitivity to the strategic nature of the audit setting.

Regardless of these potential limitations, the broader point remains. That is, as auditors focus resources away from accounts they view as low-risk, opportunities for fraud grow among those accounts. Without a strategic prompt similar to that tested here, many auditors could fail to appreciate this strategic risk, resulting in a misallocation of audit resources and more undetected misstatements.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 2002. *Consideration of Fraud in a Financial Statement Audit.* Statement on Auditing Standards No. 99. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2003. *Consideration of Fraud in a Financial Statement Audit.* Statement on Auditing Standards No. 99. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2006. *Risk Assessment Standards.* Statements on Auditing Standards Nos. 104–111. New York, NY: AICPA.

Anderson, U., and R. A. Young. 1988. Internal audit planning in an interactive environment. *Auditing: A Journal of Practice & Theory* 8 (1): 23–42.

Asare, S. K., G. M. Trompeter, A. M. Wright, and B. College. 2000. The effect of accountability and time budgets on auditors' testing strategies. *Contemporary Accounting Research* 17 (4): 539–560.

Bell, T. B., M. E. Peecher, and I. Solomon. 2005. *The 21st Century Public Company Audit. Conceptual Elements of KPMG's Global Audit Methodology.* University of Illinois at Urbana–Champaign, IL: KPMG.

Berg, J., J. Dickhaut, and K. McCabe. 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10: 122–142.

Bhattacharjee, S., M. J. Maletta, and K. K. Morena. 2007. The cascading of contrast effects on auditors' judgments in multiple client audit environments. *The Accounting Review* 82 (5): 1097–1117.

Bloomfield, R. 1995. Strategic dependence and inherent risk assessments. *The Accounting Review* 70 (1): 71–90.

Bloomfield, R. J. 1997. Strategic dependence and the assessment of fraud risk: A laboratory study. *The Accounting Review* 72 (4): 517–538.

Bowlin, K. O., J. Hales, and S. J. Kachelmeier. 2009. Experimental evidence of how prior experience as an auditor influences managers' strategic reporting decisions. *Review of Accounting Studies* 14 (1): 63–87.

Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (4): 933–945.

Caginalp, G., D. Porter, and V. Smith. 2001. Financial bubbles: Excess cash, momentum, and incomplete information. *The Journal of Psychology and Financial Markets* 2 (2): 80–99.

Camerer, C., T. H. Ho, and J. K. Chong. 2004. A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119 (3): 861–898.

Costa-Gomes, M. A., and G. Weizsäcker. 2008. Stated beliefs and play in normal-form games. *Review of Economic Studies* 75 (3): 729–762.

Croson, R. 2000. Thinking like a game theorist: Factors affecting the frequency of equilibrium play. *Journal of Economic Behavior and Organization* 41 (3): 299–314.

Fellingham, J. C., and D. P. Newman. 1985. Strategic considerations in auditing. *The Accounting Review* 60 (4): 634–650.

Fischbacher, U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10 (2): 171–178.

Fischbacher, U., and U. Stefani. 2007. Strategic errors and audit quality: An experimental investigation. *The Accounting Review* 82 (3): 679–704.

Fudenberg, D., and J. Tirole. 1991. *Game Theory.* Cambridge, MA: MIT Press.

Glover, S. M. 1997. The influence of time pressure and accountability on auditors' processing of nondiagnostic information. *Journal of Accounting Research* 35 (2): 213–226.

Haynes, C. M., and S. J. Kachelmeier. 1998. The effects of accounting contexts on accounting decisions: A synthesis of cognitive and economic perspectives in accounting experimentation. *Journal of Accounting Literature* 17: 97–136.

Hoffman, V. B., and M. F. Zimbelman. 2009. Do strategic reasoning and brainstorming help auditors change their standard audit procedures in response to fraud risk? *The Accounting Review* 84 (3): 811–837.

Houston, R. W. 1999. The effects of fee pressure and client risk on audit seniors' time budget decisions. *Auditing: A Journal of Practice & Theory* 18 (2): 70–86.

Kachelmeier, S. J., and R. R. King. 2002. Using laboratory experiments to evaluate accounting policy issues. *Accounting Horizons* 16 (3): 219–233.

Kachelmeier, S. J., and W. F. Messier, 1990. An investigation of the influence of a nonstatistical decision aid on auditor sample size decisions. *The Accounting Review* 65 (1): 209–226.

King, R. R. 2002. An experimental investigation of self-serving biases in an auditing trust game: The effect of group affiliation. *The Accounting Review* 77 (2): 265–284.

Kinney, W. R. 2005. Twenty-five years of audit deregulation and re-regulations: What does it mean for 2005 and beyond? *Auditing: A Journal of Practice & Theory* 24: 89–109.

Knechel, W. R. 2007. The business risk audit: Origins, obstacles and opportunities. *Accounting, Organizations and Society* 32 (4–5): 383–408.

Levene, H. 1957. Robust tests for equality of variances. In *Stanford Studies in Mathematics and Statistics*. Palo Alto, CA: Stanford University Press.

McNeil, K., I. Newman, and F. J. Kelley. 1996. *Testing Research Hypotheses with the General Linear Model*. Carbondale, IL: Southern Illinois University Press.

Newman, D. P., E. Patterson, and J. R. Smith. 2005. The role of auditing in investor protection. *The Accounting Review* 80 (1): 289–313.

Newman, P., S. Rhoades, and R. Smith. 1996. Allocating audit resources to detect fraud. *Review of Accounting Studies* 1 (2): 161–82.

O'Donnell, E., and J. J. Schultz. 2005. The halo effect in business risk audits: Can strategic risk assessment bias auditor judgment about accounting details? *The Accounting Review* 80 (3): 921–939.

Ochs, J. 1995. Games with unique, mixed strategy equilibria: An experimental study. *Games and Economic Behavior* 10: 202–217.

Ortmann, A., J. Fitzgerald, and C. Boeing. 2000. Trust, reciprocity, and social history: A re-examination. *Experimental Economics* 3 (1): 81–100.

Public Company Accounting and Oversight Board (PCAOB). 2007. *Report on the Second-Year Implementation of Auditing Standard No. 2: An Audit of Internal Control over Financial Reporting Performed in Conjunction with an Audit of Financial Statements*. Washington, D.C.: PCAOB.

Peecher, M. E., R. Schwartz, and I. Solomon. 2007. It's all about audit quality: Perspectives on strategic systems auditing. *Accounting, Organizations and Society* 32: 463–485.

Phillips, F. 1999. Auditor attention to and judgments of aggressive financial reporting. *Journal of Accounting Research* 37 (1): 167–189.

Rittenberg, L. E., and B. J. Schwieger. 2005. *Auditing: Concepts for a Changing Environment, 5e*. Mason, OH: Thomson South-Western.

Schuler, R. S. 1980. Definition and conceptualization of stress in organizations. *Organizational Behavior and Human Performance* 25: 184–215.

Shibano, T. 1990. Assessing audit risk from errors and irregularities. *Journal of Accounting Research* 28: 110–140.

Stahl, D., and P. A. Wilson. 1995. On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior* 10: 218–254.

Sugden, R. 1995. A theory of focal points. *The Economic Journal* 105 (430): 533–550.

Tversky, A., and D. Kahneman. 1986. Rational choice and the framing of decisions. *The Journal of Business* 59 (4): 251–278.

Weil, J. 2004. Missing numbers—Behind wave of corporate fraud: A change in how auditors work. *Wall Street Journal* (March 25): A1.

Wilks, T. J., and M. F. Zimbelman. 2004. Using game theory and strategic reasoning concepts to prevent and detect fraud. *Accounting Horizons* 18 (3): 173–184.

Zimbelman, M. F., and W. S. Waller. 1999. An experimental investigation of auditor-auditee interaction under ambiguity. *Journal of Accounting Research* 37: 135–155.

APPENDIX A

Following is the auditor's expected utility and the solution for the manager's equilibrium strategy, q_i*:

$$EU_{auditor} = \overbrace{FK}^{\text{Available resources}} - \overbrace{\sum_i e_i K}^{\text{Resources used}}$$

$$+\sum_i [\ \overbrace{Yq_ie_iT}^{\text{Detected fraud}} + \overbrace{(Y-Z)q_i(1-e_iT)}^{\text{Undetected fraud}} + \overbrace{Yr_ie_iT(1-q_i)}^{\text{Detected error}} + \overbrace{(Y-Z)r_i(1-q_i)(1-e_iT)}^{\text{Undetected error}}$$

$$+\overbrace{Y(1-r_i)(1-q_i)}^{\text{No misstatement}}]$$

$$\Rightarrow \frac{\partial EU_{auditor}}{\partial e_i} = -K + Zq_iT + Zr_iT - Zr_iq_iT = 0$$

$$\Rightarrow q_i^* = \frac{K - Zr_iT}{T(Z - Zr_i)}.$$

Following is the manager's expected utility and the solution for the auditor's equilibrium strategy, e_i*:

$$EU_{manager} = \sum_i [\overbrace{M_Hq_i(1-e_iT)}^{\text{Undetected override}} + \overbrace{(M_H - V_H)q_ie_iT}^{\text{Detected override}} + \overbrace{M_L(1-r_i)(1-q_i)}^{\text{No misstatement}} + \overbrace{M_Lr_i(1-q_i)(1-e_iT)}^{\text{Undetected Error}}$$

$$+\overbrace{(M_L - V_L)r_ie_iT(1-q_i)}^{\text{Detected Error}}$$

$$\Rightarrow \frac{\partial EU_{manager}}{\partial q_i} = M_H - V_He_iT - M_L + V_Lr_ie_iT = 0$$

$$\Rightarrow e_i^* = \frac{M_H - M_L}{T(V_H - V_Lr_i)},$$

where:

F = the total number of resource units available to the auditor;
K = the auditor's marginal utility per unit of unused resources;
Y = the auditor's payoff for a completed audit of account i;
Z = the auditor's penalty for a failed audit of account i;
T = the marginal effect per unit of resources on probability of detecting a misstatement in account i;
M_H = the manager's payoff for undetected override in account i;
M_L = the manager's payoff for no fraud or undetected error in account i;
V_H = the manager's penalty for a detected override in account i;
V_L = the manager's penalty for a detected error in account i;
e_i = the number of resource units allocated to the audit of account i;
q_i = the probability that the manager will override the account i; and
r_i = the (exogenous) probability of error in account i conditional on no override.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1255–1288

American Accounting Association
DOI: 10.2308/accr-10040

Financial Reporting Quality and Investment Efficiency of Private Firms in Emerging Markets

Feng Chen
Ole-Kristian Hope
University of Toronto

Qingyuan Li
Wuhan University

Xin Wang
The Chinese University of Hong Kong

ABSTRACT: Prior research shows that financial reporting quality (FRQ) is positively related to investment efficiency for large U.S. publicly traded companies. We examine the role of FRQ in private firms from emerging markets, a setting in which extant research suggests that FRQ would be less conducive to the mitigation of investment inefficiencies. Earlier studies show that private firms have lower FRQ, presumably because of lower market demand for public information. Prior research also shows that FRQ is lower in countries with low investor protection, bank-oriented financial systems, and stronger conformity between tax and financial reporting rules. Using firm-level data from the World Bank, our empirical evidence suggests that FRQ positively affects investment efficiency. We further find that the relation between FRQ and investment efficiency is increasing in bank financing and decreasing in incentives to minimize earnings for tax purposes. Such a connection between tax-minimization incentives and the informational role of earnings has often been asserted in the literature. We provide explicit evidence in this regard.

We have received valuable comments from Steve Kachelmeier (senior editor), Wayne Thomas, Alastair Lawrence, two anonymous reviewers, and workshop participants at The University of New South Wales, Florida International University, University of Toronto, University of Minnesota, KU Leuven, Norwegian School of Economics and Business Administration, Singapore Management University, University of Notre Dame, Santa Clara University, University of North Texas, Temple University, the University of International Business and Economics Conference (Beijing), American Accounting Association International Section Midyear Meeting (Pasadena; Best Paper Award), University of Waterloo Conference on Financial Reporting Quality in Emerging Markets, Oklahoma State University Accounting Research Conference, European Accounting Association Annual Meeting (Istanbul), Canadian Academic Accounting Association Annual Meeting (Vancouver), and American Accounting Association Annual Meeting (San Francisco). Professor Hope gratefully acknowledges the financial support of the Deloitte Professorship and Professor Li acknowledges the financial support from the National Nature Science Foundation of China (NSFC-71072103).

Editor's note: Accepted by Steven Kachelmeier.

Submitted: January 2010
Accepted: November 2010
Published Online: April 2011

Keywords: *investment efficiency; under- and overinvestment; financial reporting quality; private firms; emerging markets; financing sources; tax incentives.*

I. INTRODUCTION

This study examines the relation between financial reporting quality (FRQ) and investment efficiency for a sample of private firms in emerging markets.[1] We believe that such a study is important for the following reasons. First, and foremost, it is inherently interesting to examine settings in which the mitigation of investment inefficiencies is less likely. As discussed in detail below (see Section II), widely cited prior research concludes that private firms have lower FRQ than do public firms. The most common explanation offered is that private firms face less demand for high-quality financial information. In addition, a long line of research suggests that the value relevance of accounting information is lower in less developed countries than in more developed countries. Combining these two findings from prior research, this study examines the importance of accounting information in a setting that, *ex ante*, is expected to be less conducive to the mitigation of under- and overinvestment observed in the literature for U.S. publicly traded firms (e.g., Biddle and Hilary 2006; Hope and Thomas 2008; McNichols and Stubben 2008; Biddle et al. 2009).

Second, notwithstanding the fact that private firms (i.e., firms that are not traded on public stock exchanges) are the predominant organization in most countries, little is known about private firms' financial reporting. Compared with the large literature on developed countries' accounting systems and managers' reporting incentives, much less is known about the role of accounting in emerging markets. Research on private firms from emerging markets is virtually nonexistent despite its importance to international organizations such as the World Bank, the International Accounting Standards Board (IASB), and others.

Finally, our study complements and extends prior research on the economic consequences of variations in FRQ. In addition to investigating the overall effect of FRQ on investment efficiency, we introduce two conditional hypotheses. First, high-quality accounting information is likely more desirable in mitigating information asymmetry for private firms when they are in need of external financing. Thus, we examine whether private firms' investment efficiency is more sensitive to FRQ when these firms seek bank financing. Although prior research has examined the importance of financing sources in other settings, we are unaware of tests relating financing sources, FRQ, and investment efficiency for private firms from emerging markets. We also are unaware of studies that examine the importance of bank financing relative to other non-public equity external financing.

As our second conditional test, we consider the role of tax incentives, as prior studies generally ignore tax considerations (Hanlon and Heitzman 2010). Our study contributes by examining how tax incentives affect the strength of the relation between FRQ and investment decisions. Based on prior literature, tax considerations are especially important for private firms. In addition, there is some indication from cross-country analyses that conformity in book and tax reporting is associated with lower quality earnings (Atwood et al. 2010) and that it distorts investment decisions (Cummins et al. 1994). As book-tax conformity is higher, on average, in emerging markets than in developed countries, examining firms in emerging markets makes it possible to investigate whether the investment-FRQ relation still exists in an environment where tax considerations are important and where book-tax conformity could distort investment decisions (Cummins et al. 1994).

We obtain data from the World Bank's Enterprise Survey (WBES), a major cross-sectional firm survey conducted around the world by the World Bank. This database has been used in a number of prior studies (e.g., Beck and Demirguc-Kunt 2006; Beck et al. 2005, 2006, 2008; Carlin et al. 2007; Brown et al. 2008; Hope et al. 2011).[2] Unlike studies of the role of accounting information in influencing investment decisions that only use data on public firms, WBES data allow us to examine small- and medium-size private enterprises. Further, the WBES database includes information on both the sources of financing (e.g., bank financing, internal financing, informal financing) used in making new investments and the degree to which the firm faces income tax pressures (i.e., higher tax rates and stronger tax enforcement), without having to rely on proxies. This unique database allows us to investigate whether the effects of FRQ on investment efficiency varies with the source of financing and with firms' tax incentives.

---

[1] Consistent with Biddle et al. (2009), we define FRQ as the precision with which financial reporting conveys information about firms' operations.

To generalize our results and reduce measurement error, we use several proxies for FRQ in our empirical tests. Specifically, we use (1) the Kothari et al. (2005) discretionary accruals measure; (2) the McNichols and Stubben (2008) and Stubben (2010) revenue-based measure; (3) the Dechow and Dichev (2002) measure as implemented by Francis et al. (2005) and Srinidhi and Gul (2007); and (4) a summary statistic formed by aggregating these three measures. Analyzing WBES data from 2002 to 2005, our main findings are as follows. First, all four proxies for FRQ are statistically and economically significantly related to investment efficiency. Specifically, all four of our FRQ proxies are significantly negatively associated with both under- and overinvestment. Second, the importance of FRQ is increasing in the degree of bank financing, likely reflecting the use of financial statements by banks in granting credit. Third, for firms facing greater income tax pressures, the relation between FRQ and investment efficiency is reduced. It is often argued in the literature that a focus on minimizing taxes can negatively affect the usefulness of financial statement information; however, to date, evidence on this issue has been limited. Our results are robust to the inclusion of firm-level control variables, industry and country fixed effects, to the use of alternative investment models to measure investment efficiency (including a change specification), to alternative measures of FRQ, to endogeneity, and to several other controls.

Our study advances the literature by providing empirical evidence that FRQ enhances firm-level investment efficiency in a sample of small, private firms across 21 emerging markets. Given that prior widely cited research suggests that FRQ should be considerably lower for these firms than for publicly traded companies in developed countries, this study can be viewed as an examination of "boundary conditions" for the importance of accounting information. Our findings concerning how the importance of FRQ varies with financing sources and tax incentives complement and extend current research on the relation between FRQ and investment efficiency (e.g., Biddle and Hilary 2006; McNichols and Stubben 2008; Biddle et al. 2009). Existing research examines only listed firms, so that even the smallest firms in prior studies are relatively large. In contrast, we focus on a sample of private firms that are important drivers of economic growth globally and for which there is limited extant research.[3]

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

One objective of financial reporting information is to facilitate the efficient allocation of capital. An important aspect of this role is to improve firms' investment decisions. Specifically, theory suggests that improved financial transparency has the potential to alleviate both under- and overinvestment problems and recent studies support this prediction (e.g., Biddle and Hilary 2006;

---

[2] For a more extensive list, please see http://www.enterprisesurveys.org/documents/Research-used-in-different-studies.xls.

[3] Guidance on how to improve the quality of investment decisions in emerging markets is directly relevant to the World Bank's mission of "offering assistance to developing countries around the world." See http://www.enterprisesurveys.org.

Hope and Thomas 2008; McNichols and Stubben 2008; Biddle al. 2009). This evidence, however, has been mostly limited to large, publicly traded companies in the United States.

We examine the role of financial reporting quality in a very different setting from that examined in prior studies. Specifically, we deliberately "turn two dials at once" by moving from public to private firms and by moving from developed countries to emerging markets—allowing us to examine the importance of FRQ under conditions that prior research suggests are less conducive to the mitigation of under- and overinvestment previously observed for publicly traded enterprises in the United States. As such, our study can be viewed as a natural laboratory for examining boundary conditions for the relevance of FRQ for investment efficiency in several ways.

First, in contrast to prior research, we focus on *private* firms. More than 99 percent of limited liability companies, in most countries, are not listed on a stock exchange (e.g., Pacter 2004; Berzins et al. 2008; Nagar et al. 2011). In the aggregate, non-listed firms have about four times more employees, three times higher revenues, and twice the amount of assets than do listed firms (Berzins et al. 2008). In spite of their economic importance and likely differences from public companies, comparatively little is known about financial reporting of private firms.

Private firms are different from publicly traded firms in several respects. Private firms are more closely held and have greater managerial ownership. Moreover, their major capital providers often have insider access to corporate information and typically take a more active role in management (e.g., Van Tendeloo and Vanstraelen 2008). With greater ownership concentration, large shareholders can take advantage of their controlling positions and direct private benefits for personal consumption, which is the typical problem of expropriation of minority shareholders and creditors (e.g., Morck et al. 1988). Furthermore, given the stronger ownership concentration, shareholder turnover is lower, and shareholders take a more active role in management, which reduces their reliance on financial statements for monitoring managers compared with public firms (Ball and Shivakumar 2005). Finally, private firms' financial statements are not as widely distributed to the public and are more likely to be influenced by taxation, dividend, and other objectives (Ball and Shivakumar 2005).

Prior studies find evidence suggesting that private firms have lower earnings quality on average than do public firms. For example, Ball and Shivakumar (2005) show that private U.K. companies exhibit less timely loss recognition than do public companies. Using a European dataset, Burgstahler et al. (2006) find that private firms exhibit lower quality earnings. Both studies argue that the main explanation for their findings of lower FRQ is the lower market demand for high-quality financial reporting for such firms.

Thus, based on prior research, there are reasons to believe that FRQ is lower for private firms than for public firms.[4] Consequently, it is not clear whether results from U.S. publicly traded companies will hold for our sample of small, private firms.

Second, in addition to focusing on private firms and in contrast to prior research, we also focus on firms from emerging markets. An extensive line of research exists that provides evidence of variation in accounting standards, accounting practices, enforcement of accounting rules, properties of accounting earnings, and the value relevance of accounting information around the world (e.g.,

---

[4] Not all extant research provides this conclusion. In a recent study, Givoly et al. (2010) find that whereas U.S. firms with privately held equity have less timely loss recognition, they have higher quality accruals and lower propensity to manage income *vis-à-vis* earnings thresholds than do private equity firms. However, they have a rather small sample of firms classified as "private" (531) and these firms all issue public debt, and are thus really a hybrid form of companies (and classified as "public" by the Securities and Exchange Commission). In addition, the sample firms are owned by financial sponsors and/or management. Finally, the sample firms are very large compared to the typical private firms and almost all employ a Big 4 auditor (94 percent). Thus, their results may not generalize to the typical U.S. private firms. In fact, we view further exploration of FRQ in private versus public firms as fertile ground for future research.

Alford et al. 1993; Ali and Hwang 2000; Ball et al. 2000; Hung 2000; Hope 2003; Leuz et al. 2003; Hail and Leuz 2006). For example, Ball et al. (2000) argue that the role of accounting information is more limited in environments that are characterized by low investor protection and more concentrated ownership structures. Ali and Hwang (2000) show that the value relevance of accounting information is lower for countries with bank-oriented (as opposed to market-oriented) financial systems and for countries with a greater degree of conformity between financial accounting and tax rules. Similarly, Atwood et al. (2010) show that earnings have lower persistence and a lower association with future cash flows when book-tax conformity is higher. The emerging markets included in our sample can, on average, be characterized as having lower investor protection, greater ownership concentration, higher book-tax conformity, and more bank-based financing than do the United States and other highly developed markets. In sum, our tests using private firms from emerging markets can be viewed as a test of the "boundary condition" (or lower bound) for the importance of FRQ in ameliorating investment efficiencies.

While there are clear reasons to expect that FRQ will play a less prominent role for our sample firms, counter-arguments do exist. The economic theories that provide a role for FRQ are not limited to public firms, although the effect may be magnified in a public-firm setting. There are several mechanisms through which financial reporting can mitigate under- and overinvestment problems. First, accounting information can aid investment efficiency by reducing adverse selection, liquidity risk, and information risk (Diamond and Verrecchia 1991; Leuz and Verrecchia 2000; Easley and O'Hara 2004; Lambert et al. 2007). In the absence of public equity markets, these effects could be muted, but could still exist for non-equity financing. Second, disclosed financial information aids corporate control mechanisms in preventing managers from expropriating wealth from investors or creditors (e.g., Fama and Jensen 1983), hence, providing external suppliers of capital with greater assurance about managers' activities as well as aiding in internal stewardship functions (e.g., board supervision of management). Third, improved accounting information can enhance the efficiency with which managers make investment decisions.[5]

In addition, private firms typically have a weaker information disclosure environment than that characterizing public firms (e.g., Burgstahler et al. 2006). This observation suggests that, even if FRQ is lower for private firms, accounting information could still play a role because there are fewer competing sources of information.[6] As noted above, McNichols and Stubben (2008) emphasize the role that accounting information plays in internal decision making. Small firms are unlikely to have management accounting systems that are separate from financial accounting (e.g., Drury and Tayles 1995), potentially enhancing the role of accounting in internal decision making.

Finally, it is possible that the lack of analyst coverage, lower media coverage, and overall lower-quality institutions in emerging markets makes accounting information a relatively greater component of the overall information set used for decision making by insiders or outsiders.[7,8]

---

[5] As McNichols and Stubben (2008, 1571) point out, investment decisions depend on expectations of investment benefits. These benefits in turn depend on expectations of future growth and product demand. In other words, high-quality information can help managers form more accurate expectations and identify better investment opportunities, thereby improving investment efficiency even in a world without adverse selection and/or moral hazard (Bushman and Smith 2001; McNichols and Stubben 2008).

[6] For example, Indjejikian and Matejka (2009) highlight the importance of accounting information for private firms in compensation contracts.

[7] Abu-Nassar and Rutherford (1996), Mirshekary and Saudagaran (2005), and Al-Razeen and Karbhari (2007) conduct surveys and conclude that accounting information is the most important information source for users in emerging markets.

[8] Accounting information may be useful for bank financing, for attracting new equity capital (for the private firms with more dispersed ownership and/or firms that are selling shares to new investors), to suppliers in their decisions to grant trade credit, to other providers of finance such as leasing companies, and to non-manager employees (e.g., Bova et al. 2011).

Based on the above discussion, we follow the approach of Biddle et al. (2009) and test whether the FRQ of private firms from emerging markets helps mitigate both under- and overinvestment. In other words, we investigate if FRQ also mitigates capital investment inefficiencies under less conducive conditions than those examined in prior research. Our first hypothesis is as follows:

H1: Financial reporting quality mitigates both underinvestment and overinvestment.

In addition to examining the overall effect of FRQ on investment efficiency, we explore two conditional effects: financing sources and tax incentives. We first investigate the effect of financing sources on the effect of FRQ on investment efficiency. WBES data provide a detailed breakdown of financing sources as a percentage of new investment. Private firms in emerging markets fund investment from external sources including bank lending, issuance of private equity, leasing, trade credit, financing from special development agencies or governments, and informal financing (i.e., financing from moneylenders, family, and friends), and from retained earnings and additional contributions by owners (Beck et al. 2008). A large body of literature documents how the availability of external or internal funds affects investment decisions (Myers and Majluf 1984; Fazzari et al. 1988; Blanchard et al. 1994), where these prior studies have used samples of public firms that rely mainly on equity and debt financing. In contrast, for private firms, external financing sources are usually limited and consist mainly of bank loans and trade credit.

We argue that, the financing source will affect the relation between FRQ and investment efficiency. Specifically, we are interested in whether firms that rely more on bank financing have a stronger relation between FRQ and investment efficiency than do other private firms in emerging economies.

Bank lending is the most common source of external capital for private firms in developing countries (Beck et al. 2008; Brown et al. 2008). Banks may well have access to additional information beyond the financial statements, potentially reducing the importance of accounting information. However, besides the large body of research documenting the role of accounting information for lending decisions in the U.S. and other developed countries, there is also extensive evidence that banks rely on borrowers' financial reports in credit decisions in emerging markets and for small firms (e.g., Danos et al. 1989; Berry et al. 1993; Berry et al. 2004; Kitindi et al. 2007). Compared with other external capital suppliers that rely more on mutual trust and private communication, banks are likely to screen the financial statements of corporate clients more carefully. Importantly, banks not only lend larger amounts compared with other sources of informal financing, but also extend loans with longer maturities, which makes them more vulnerable to information and incentive problems. Failure rates are higher among smaller firms, further encouraging banks to carefully examine financial information in their lending decisions. Examining a client's financial statements helps banks to determine the firm's assets that can serve as collateral, to evaluate its future cash-flow generating capability, to gauge the firm's debt capacity, and to analyze the riskiness of the firm in determining a lending rate. In other words, the importance of FRQ should increase with the extent of bank financing compared with other, more informal sources of financing. This discussion motivates our second hypothesis:[9]

H2: The relation between financial reporting quality and investment efficiency is stronger if a firm's investment is mainly funded through bank financing.

Finally, we explore the conditional effect of tax incentives. We know from prior research that private firms are influenced relatively more by tax objectives than are public firms (e.g., Ball and Shivakumar 2005). Further, the alignment between financial and tax accounting is higher for our sample than for publicly traded firms in developed countries. High alignment means that financial statements serve as the basis for taxation or that tax laws explicitly require an equivalent treatment for certain items in both sets of accounts (e.g., Burgstahler et al. 2006). In such an environment, a focus on minimizing taxes can directly impact FRQ, and we test for the effect of cross-sectional variation in tax incentives. That is, for firms that face especially high income tax rates and strong enforcement by tax authorities (i.e., high "tax pressure"), the primary objective of financial reporting could be to minimize income taxes rather than to provide information to suppliers of capital or to management, which could reduce the role of FRQ on investment efficiency. This reasoning motivates our final hypothesis:

---

[9] Note that our hypothesis differs from Biddle and Hilary (2006), who compare bank financing with public equity financing across countries. In contrast, as private firms do not rely on public equity financing, we compare the role of bank financing primarily with other non-public debt, private equity, and informal financing.

> **H3:** The relation between financial reporting quality and investment efficiency is less pronounced for firms that have strong incentives to manage their earnings for tax purposes.

III. SAMPLE AND MEASUREMENT OF MAIN VARIABLES

Data Source

We obtain our data from World Bank's Enterprise Survey (WBES) conducted during 2002–2005 by the World Bank in 79 countries, including many low-income countries.[10] Prior studies that have used this database include Beck and Demirguc-Kunt (2006), Beck et al. (2005, 2006, 2008), Carlin et al. (2007), Brown et al. (2008), and Hope et al. (2011).

The primary goal of WBES is to provide quantitative data that allow an assessment of a country's investment climate in an internationally comparable manner.[11] The surveys are administered in face-to-face interviews with managing directors, accountants, human resource managers, and other relevant company staff.[12] Samples are stratified by size, industry, and location. Although the surveys are conducted with the knowledge and support of relevant government authorities, governments are not provided with the raw data or other information that would allow them to identify the responses of individual firms, and businesses are informed of this confidentiality prior to interviews to encourage truthful responses.[13]

The dataset includes a large sample of firms across multiple sectors (manufacturing, services, agriculture, and construction). Usable data include both quantitative and qualitative information on firm characteristics, including sources of finance, barriers to growth, access to infrastructure services, legal difficulties, and corruption. The dataset also includes measures of firm performance, such as multiple years of historical data on capital investment and operating performance.

A limitation of using survey data is that financial statement information is necessarily restricted. Thus, we do not have access to some potential control variables used in prior studies. However, the survey does contain data on a number of important firm characteristics and, consistent with prior research that employs the WBES data, we include these firm characteristics as well as industry and country fixed effects in our empirical tests.

---

[10] Although WBES also has some data before and after our sample period, the availability of accounting information is very limited in the other years.

[11] This discussion draws on *Productivity and Investment Climate Survey (PICS): Implementation Manual* (IBRD 2003).

[12] WBES begins from a minimum core set of questions that are common across all countries. However, survey managers at the country or region level are allowed to extend the survey. To maintain cross-country comparability, core questions cannot be reworded except in translating to well-understood phrases with the same meaning. In addition, all core questions are asked using standardized instructions provided by the survey's designers. Finally, note that since the data are based on face-to-face interviews, the World Bank data do not suffer from the typical issues related to survey response bias as when using mail- or web-based surveys.

[13] Carlin et al. (2007), Beck et al. (2008), and Hope et al. (2011) contain several validity tests of the WBES data.

We describe the sample selection procedure in Table 1. The main constraints on our sample size are the availability of data to compute the investment efficiency measure and the four FRQ measures. Each firm appears only once in the database (and hence the number of observations equals the number of unique firms). However, firms are requested to not only provide data for the current year, but also for the previous two years. The requirement to have previous years' data (necessary for our tests) is the main constraint on our sample size. We utilize 6,727 unique firms for the test of H1 involving discretionary revenues (described below) as our FRQ proxy, and somewhat smaller sample sizes for tests using the other three FRQ proxies. The sample size is reduced further for the tests of H2 and H3 for which we require additional data for the firm's financing sources and tax burden.

Table 2 provides a distribution of the sample by country. A total of 21 countries are represented, with the greatest number of observations coming from Thailand, Brazil, Pakistan, Vietnam, and India. These five countries make up 63 percent of the total sample. Although all of our sample countries are considered "developing" or emerging markets as per United Nations classifications, there is considerable variation across our sample countries. We examine the effects of FRQ on investment efficiency across country characteristics in Section V.

Proxy for Investment Efficiency

The two key constructs in the analysis are investment efficiency and financial reporting quality, and we investigate how FRQ in the current year affects *next* year's investment efficiency.

TABLE 1

Sample Selection

All firm-years in the WBES database with listing status information	47,712
Less firms residing in "Other" sectors	770
Less publicly traded firms	3,200
Private firms with all necessary industry information	43,742
Less firms without information in the financial statement section	18,658
Less firms with missing data on revenue growth in the prior year	9,146
Less firms with missing data on investment	5,869
Additional deduction by data requirement of investment model (i.e., at least ten firms in each industry by country for Equation (1))	77
Firms with the investment efficiency variable (*InvEff*)	9,992
(1) Missing data on discretionary revenue (*DisRev*) and control variables (Table 4)	3,265
Sample size for the main tests with *DisRev* as a proxy for financial reporting quality	6,727
(2) Missing data on discretionary accruals (*DisAccr*) and control variables (Table 4)	3,520
Sample size for the main tests with *DisAccr* as a proxy for financial reporting quality	6,472
(3) Missing data on the modified Dechow-Dichev measure (*DD*) and control variables (Table 4)	3,604
Sample size for the main tests with *DD* as a proxy for financial reporting quality	6,388
(4) Firms with missing data for any of the three financial reporting quality measures (*DisRev*, *DisAccr*, or *DD*) or for control variables.	3,671
Sample size for the main tests with aggregate FRQ (*Aggreg*) as a proxy for financial reporting quality	6,321

TABLE 2
Sample Distribution by Country

Country	n	Percentage
Bangladesh	483	7.17%
Brazil	1,033	15.35%
Ecuador	55	0.81%
El Salvador	88	1.31%
Eritrea	24	0.36%
Ethiopia	36	0.54%
Guatemala	65	0.97%
Honduras	41	0.61%
India	486	7.22%
Indonesia	314	4.67%
Nicaragua	39	0.58%
Oman	36	0.54%
Pakistan	850	12.64%
Philippines	371	5.52%
South Africa	388	5.77%
Sri Lanka	296	4.40%
Syria	44	0.65%
Tanzania	75	1.11%
Thailand	1,307	19.43%
Vietnam	587	8.73%
Zambia	109	1.62%
Total	6,727	100%

This table shows the sample distribution by country. There are a total number of 6,727 observations from 21 countries.

Conceptually, investment efficiency refers to firms undertaking all and only projects with positive net present value. Consistent with prior research (e.g., Biddle et al. 2009), we measure investment efficiency as deviations from expected investment using a model that predicts investment as a function of growth opportunities. Thus, both underinvestment (negative deviations from expected investment) and overinvestment (positive deviations from expected investment) are considered inefficient investments. Specifically, we estimate a parsimonious model for expected investment as a function of revenue growth (see, e.g., Modigliani and Miller 1958; Hubbard 1998). As the relation between investment and revenue growth could differ between revenue decreases and revenue increases (e.g., Eberly 1997; McNichols and Stubben 2008), we allow for differential predictability for revenue increases and revenue decreases by employing a piecewise linear regression model:

$$Invest_{i,t} = \alpha_0 + \alpha_1 NEG_{i,t-1} + \alpha_2 \%RevGrowth_{i,t-1} + \alpha_3 NEG * \%RevGrowth_{i,t-1} + \varepsilon_{i,t}. \qquad (1)$$

Following Biddle et al. (2009), we define $Invest_{i,t}$ as the sum of new investment in machinery, equipment, vehicles, land, buildings, and research and development expenditures, less the sale of fixed assets, and scaled by lagged total assets for firm i in year t. $\%RevGrowth_{i,t-1}$ is the annual revenue growth rate for firm i in year t−1. The indicator variable $NEG_{i,t-1}$ takes the value of 1 for negative revenue growth, and 0 otherwise.

We estimate the investment model cross-sectionally with at least ten observations in each WBES industry by country. The sample consists of 9,992 firm-year observations with available data to estimate Equation (1). To mitigate the influence of outliers, we winsorize all variables at the 1 percent and 99 percent levels. We then classify firms into two groups based on the residuals of Equation (1) (i.e., the deviations from the predicted investment levels). To ease exposition, we multiply the underinvestment variable by −1 so that a higher value suggests a more severe underinvestment.[14]

Proxies for Financial Reporting Quality

There is no universally accepted measure of FRQ. We employ three measures that have been used in prior research as well as an aggregate measure for the following reasons. First, a single proxy is unlikely to cover all facets of FRQ. Second, the use of multiple proxies helps to generalize our results. Third, using alternative measures mitigates the possibility that results using one particular proxy capture some factor other than FRQ, and that this other factor is driving our results.

The first measure is performance-adjusted discretionary accruals as developed by Kothari et al. (2005). Specifically, we estimate the following model by country and for each industry that has at least 16 observations:

$$TAccr_{i,t} = \alpha_0 + \alpha_1(1/Assets_{i,t-1}) + \alpha_2\Delta Rev_{i,t} + \alpha_3 PPE_{i,t} + \alpha_4 ROA_{i,t} + \varepsilon_{i,t}, \qquad (2)$$

where $TAccr_{i,t}$ is total accruals, measured as the change in non-cash current assets minus the change in current non-interest-bearing liabilities, minus depreciation and amortization expense for firm *i* at year *t*, scaled by lagged total assets ($Assets_{i,t-1}$); $\Delta Rev_{i,t}$ is the annual change in revenues scaled by lagged total assets; $PPE_{i,t}$ is property, plant, and equipment for firm *i* at year *t*, scaled by lagged total assets; $ROA_{i,t}$ is return on assets for firm *i* at year *t*. The residuals from the regression model are discretionary accruals. In our tests, we use the absolute values of discretionary accruals as a proxy for FRQ. We multiply the absolute values of discretionary accruals by −1 (*DisAccr*). Thus, higher values of *DisAccr* represent higher FRQ.

To calculate the second proxy, we follow McNichols and Stubben (2008) and Stubben (2010) and estimate discretionary revenues. Specifically, we use the following regression:

$$\Delta AR_{i,t} = \alpha_0 + \alpha_1\Delta Rev_{i,t} + \varepsilon_{i,t}, \qquad (3)$$

where $\Delta AR_{i,t}$ represents the annual change in accounts receivable and $\Delta Rev_{i,t}$ is the annual change in revenues, each scaled by lagged total assets. Discretionary revenues are the residuals from Equation (3), which is estimated separately for each industry-country group that has at least eight observations. We multiply the absolute values of discretionary revenues by −1 (*DisRev*). Thus, higher values of *DisRev* represent higher FRQ.

Our third proxy is based on the cross-sectional Dechow and Dichev (2002) model, as modified by McNichols (2002) and Francis et al. (2005). Specifically, we estimate the following model by country and for each industry that has at least 16 observations:

$$TCAccr_{i,t} = \alpha_0 + \alpha_1 OCF_{i,t-1} + \alpha_2 OCF_{i,t} + \alpha_3 OCF_{i,t+1} + \alpha_4\Delta Rev_{i,t} + \alpha_5 PPE_{i,t} + \varepsilon_{i,t} \qquad (4)$$

where $TCAccr_{i,t}$ is total current accruals, measured as the change in non-cash current assets minus the change in current non-interest-bearing liabilities, scaled by lagged total assets; *OCF* is cash flow from operations, measured as the sum of net income, depreciation and amortization, and changes in current liabilities, minus changes in current assets, scaled by lagged total assets;[15] $\Delta Rev_{i,t}$ is the annual change in revenues scaled by lagged total assets; $PPE_{i,t}$ is property, plant, and equipment, scaled by lagged total assets. The residuals from Equation (4) represent the estimation errors in the current accruals that are not associated with operating cash flows and that cannot be explained by the change in revenue and the level of *PPE*. Given the short longitudinal time frame in our study, we follow Srinidhi and Gul (2007) and use the absolute value of this residual as a proxy for FRQ. We multiply the absolute values of the Dechow-Dichev measure by −1 (*DD*). Thus, higher values of *DD* represent higher FRQ.[16]

---

[14] The average adjusted R^2 for the investment model is 12.2 percent. To provide some benchmarks we note the following. First, when we estimate the same model for all Compustat firms (other than finance, utilities, and services industries) for 2002–2006, we obtain an R^2 of 10.1 percent. Second, McNichols and Stubben (2008) report an adjusted R^2 of 13 percent (see their Table 4, Panel A, which includes both Q and cash flows). We further note that the investment model with seven additional explanatory variables and the change specification (described later) obtain adjusted R^2s of 28.8 percent and 34.3 percent, respectively.

Last, to mitigate measurement error in the individual FRQ components and to provide evidence based on an overall FRQ metric, we aggregate the three proxies into one aggregate score. Specifically, following Biddle et al. (2009), we first normalize all proxies and then take the average of the three measures as our summary FRQ statistic (*Aggreg*).

Bank Financing and Tax Burden Variables

To test H2, we use an indicator variable, *Bank*, to capture bank financing as a major external financing source for firms. Specifically, this indicator takes a value of 1 if bank financing is the dominant source of external financing. Prior studies document that internal sources of finance predominate (e.g., Beck et al. 2008; Brown et al. 2008). To ensure that bank financing is economically important, we further require that bank financing constitutes more than 5 percent of all sources that fund new investments.[17]

Our measure of income tax pressure refers to managers' perceptions of whether tax rates and tax administration impose a major or severe obstacle in the operation and growth of their firms. In the WBES database, higher ratings of tax rates and/or tax administration suggest a heavier tax burden borne by the firm, implying that firms have greater incentives to manage earnings for tax purposes. We use an indicator variable, *Tax*, to test H3. The indicator takes a value of 1 when the average rating of managers' responses is 3 or greater.[18]

Descriptive Statistics

Table 3 provides descriptive statistics for our measures of investment efficiency, FRQ, bank financing, tax burden, as well as for our main control variables. Panel A shows that only 27 percent of firms belong to the overinvestment group, while an overriding majority of sample firms belong to the underinvestment group. This intuitive result confirms that private firms in emerging markets, due to their difficulty in securing external financing, are more likely subject to the problem of underinvestment rather than overinvestment.

For the 4,590 firms that provide information about their financing sources, 38 percent have bank financing as their dominant source of external financing (constituting more than 5 percent of

---

[15] Due to data limitations, for year $t-1$ we proxy for operating cash flows using net income adjusted for depreciation and amortization.

[16] In untabulated analyses we find that the use of the *signed* measures of three FRQ proxies (i.e., *DisAccr*, *DisRev*, and *DD*) corroborates the conclusions reached using residuals with absolute values.

[17] As sensitivity tests we (1) remove the 5 percent threshold requirement; (2) change the threshold to 10 percent; and (3) instead use an indicator value that takes the value of 1 if the firm has non-zero bank financing (0 otherwise). In all three cases inferences are unaffected. Note that bank financing includes financing from both local and foreign banks. However, only 100 sample firms have foreign bank financing, and deleting these observations does not affect any inferences.

[18] Specifically, in the WBES database, a rating of 0 denotes no obstacle; 1 denotes a minor obstacle; 2, a moderate obstacle; 3, a major obstacle; and 4, a very severe obstacle.

TABLE 3
Descriptive Statistics of the Main Sample

Panel A: Descriptive Statistics

Variable	n	Mean	25%	Median	75%	STD
InvEff (signed values)	6,727	0.002	−0.050	−0.023	0.004	0.097
Overinvestment	1,810	0.122	0.025	0.081	0.223	0.112
Underinvestment	4,917	0.043	0.018	0.037	0.061	0.030
DisRev	6,727	−0.047	−0.065	−0.024	−0.008	0.056
DisAccr	6,472	−0.100	−0.136	−0.056	−0.020	0.112
DD	6,388	−0.090	−0.117	−0.050	−0.019	0.107
SDisAccr (signed values)	6,472	−0.001	−0.063	−0.004	0.048	0.151
SDisRev (signed values)	6,727	0.000	−0.029	−0.005	0.018	0.073
SDD (signed values)	6,388	0.000	−0.058	−0.006	0.042	0.141
Aggreg	6,321	0.007	−0.335	0.296	0.611	0.820
Bank	4,590	0.382	0.000	0.000	1.000	0.486
Tax	6,695	0.174	0.000	0.000	0.000	0.379
Assets (in thousands)	6,727	51,300	172	943	6,341	181,000
Age (in years)	6,727	17.334	9.000	15.000	31.000	14.337
Tang	6,727	0.375	0.151	0.348	0.563	0.254
Slack	6,727	0.069	0.007	0.028	0.089	0.103
Audit	6,727	0.522	0.000	1.000	1.000	0.500

Panel B: Correlation Matrix

	InvEff	*DisRev*	*DisAccr*	*DD*	*Aggreg*	*Bank*	*Tax*	*LogAssets*	*LogAge*	*Tang*	*Slack*
DisRev	−0.231***										
DisAccr	−0.150***	0.417***									
DD	−0.146***	0.379***	0.836***								
Aggreg	−0.210***	0.714***	0.900***	0.884***							
Bank	−0.046***	−0.036**	−0.052***	−0.020*	−0.045***						
Tax	0.011***	0.009	0.010	0.020	0.022*	−0.087***					
LogAssets	−0.247***	0.076***	0.143***	0.140***	0.147***	0.086***	0.070***				

(continued on next page)

TABLE 3 (continued)

	InvEff	*DisRev*	*DisAccr*	*DD*	*Aggreg*	*Bank*	*Tax*	*LogAssets*	*LogAge*	*Tang*	*Slack*
LogAge	−0.064***	0.074***	0.068***	0.074***	0.089***	−0.039***	0.019	0.145***			
Tang	−0.040***	0.176***	0.205***	0.206***	0.236***	0.094***	−0.005	0.076***	−0.131***		
Slack	0.051***	−0.017	−0.042***	−0.061***	−0.043***	−0.102***	−0.025**	−0.175***	−0.043***	−0.136***	
Audit	−0.080***	0.024**	0.059***	0.078***	0.062***	0.156***	−0.088***	0.208***	0.107***	−0.036***	−0.058***

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, using two-tailed tests.

Panel A presents descriptive statistics for the variables used in the analyses. Panel B presents Pearson correlations for these variables. Variable computations are described in Appendix A unless noted here: Panel A: *InvEff* is the signed residual values from the investment efficiency model. *SDisAccr*, *SDisRev*, and *SDisRev* refer to the signed residuals from the discretionary accruals model, the discretionary revenue model, and the modified Dechow-Dichev model, respectively. *Assets* are reported in thousand dollars and *Age* is reported in years; Panel B: *DisAccr*, *DisRev*, and *DisRev* refer to the absolute value of residuals from the estimation models, multiplied by −1.

all sources that fund new investments). For the 6,695 firms that provide information about the firm-level tax burden, about 17 percent report facing severe tax obstacles. The average firm has been in existence for 17 years. As expected, there is wide variation in firm size measured by total assets. But most firms are quite small, with half the firms having total assets of $943,000 or less. In all tests we use log transformations of both firm age and firm size.

Table 3, Panel B reports Pearson correlations. As predicted, all four proxies of FRQ are significantly negatively correlated with the proxy for investment inefficiency. In addition, the FRQ proxies are positively and significantly correlated. However, as the correlation coefficients are below 1, they still capture somewhat different dimensions of FRQ, and using all four proxies in our tests increases the generalizability of our inferences.[19] As correlation results do not control for differences in firm, industry, or country characteristics, we now turn to multivariate tests.

IV. RESEARCH DESIGN AND RESULTS

Basic Empirical Model

Because we are interested in how financial reporting quality affects investment efficiency, we test H1 by regressing the measure of investment efficiency in year *t*+1 on the measures of FRQ in year *t*. Similar to Biddle et al. (2009), we also estimate Equation (5) separately for underinvestment and overinvestment. This design allows us to test whether higher FRQ mitigates both. The basic model is:

$$\begin{aligned} InvEff_{i,t+1} = {} & \beta_0 + \beta_1 FRQ_{i,t} + \beta_n Control\ Variables_{i,t} \\ & + \sum Industries + \sum Countries + \varepsilon_{i,t}, \end{aligned} \tag{5}$$

where:

InvEff = excess investment (*underinvestment* or *overinvestment*) is the residual of the investment model as described above. We use the absolute value of the residuals for *underinvestment*; and

FRQ = financial reporting quality measured in the following four ways: (1) *DisAccr*, which is the absolute residual of the Kothari et al. (2005) performance-adjusted discretionary accruals model, multiplied by −1; (2) *DisRev*, which is the absolute residual of the McNichols and Stubben (2008) discretionary revenue model, multiplied by −1; (3) *DD*, which is the absolute residual of the modified Dechow and Dichev (2002) model, multiplied by −1; and (4) *Aggreg*, which is the average of the standardized previous three measures.

Motivated by prior research, we include firm size, firm age, asset tangibility, financial slack, and external auditing as control variables, as well as industry fixed effects.[20] We also include country fixed effects in all models, which is a common approach for controlling country-specific effects and addressing correlated omitted country-level variable problems (Doidge et al. 2007).

To the extent that FRQ mitigates under- and overinvestment (H1), β_1 is expected to be negative.

---

[19] We also note that, as expected, firms that have their financial statements reviewed by auditors have higher financial reporting quality, as evidenced by the positive correlation between *Audit* and all four FRQ proxies.

[20] Untabulated tests show that our results are almost identical using number of employees or revenues as alternative proxies of firm size. Note further that we later report results of tests that include ten additional control variables. Please see Appendix A for all variable definitions.

Results

Table 4 reports the regression results for the test of H1 using all four FRQ proxies. The models have adjusted R^2s of between 14 percent and 46 percent, with a higher explanatory power for underinvestment (the most prevalent scenario in our sample). The first four columns use underinvestment as the dependent variable and the last four columns use overinvestment as the dependent variable.

Across all eight test specifications, the conclusion is the same: FRQ enhances investment efficiency. Specifically, all eight estimated FRQ coefficients are negative and significant at the 1 percent level. These results are consistent with the Pearson correlation coefficients and are robust to controls for five firm-level characteristics as well as country and industry fixed effects. Of the control variables, firm size is negatively and significantly associated with both under- and overinvestment, consistent with expectations and prior research. Asset tangibility is positively correlated with deviations from expected investment. External auditing has a mostly negative estimated coefficient but is not statistically significant.

The test results are also economically significant. For example, in the overinvestment scenario, a one standard deviation decrease in the *DisAccr* measure implies a decrease of overinvestment by 1.44 percent of total assets.

Taken together, the results in Table 4 suggest that, although prior research suggests that FRQ should be lower among private firms than among public firms and lower in emerging markets than in developed markets, we observe evidence that the quality of financial reporting affects subsequent capital investment efficiency in accordance with theory, even for private firms from emerging markets. Specifically, our results support H1 and suggest that higher FRQ enhances investment efficiency even in what prior research would classify as a "boundary condition." We present results of numerous additional analyses that test the robustness of the findings in Section V.

In addition to complementing and extending prior academic research, our findings should be relevant to the World Bank, whose mission is to aid in improving living conditions in developing countries, as it is likely that more efficient investments should lead to higher social welfare. We further believe that the findings should be important to the IASB, which is currently working on standards for small- and medium-sized enterprises (SMEs), including private companies around the world. In fact, it is a reasonable conjecture that, compared with publicly traded companies in the U.S. and similar countries, our sample firms face fewer mandatory accounting recognition and measurement (as well as disclosure) requirements. Therefore, our sample firms' choices regarding FRQ have the potential to be especially important for improving investment efficiency.

For tests of the relative role of bank financing (H2) and the effect of tax incentives (H3), we add the main effects of these variables as well as interaction effects with the FRQ measures. We predict a negative coefficient on *FRQ* × *Bank* and a positive coefficient on *FRQ* × *Tax*.

Table 5 presents results for the conditional effect of bank financing. The additional data requirement for financing sources results in a reduced sample size, ranging from 4,590 firms for the tests employing discretionary revenues to 4,234 firms for the tests using the aggregate measure. We first note that, after controlling for the effect of bank financing, all *FRQ* measures (for both under- and overinvestment) continue to load negatively and significantly (at the 5 percent level or better). The main effect of *Bank* is not significant. Our focus, however, is on *FRQ* × *Bank*, and we observe negative estimated coefficients using all four proxies and for both underinvestment and overinvestment. However, whereas all interaction effects are significant for underinvestment, statistical significance is lacking for overinvestment (possibly due to the smaller sample size available for the latter test). The results suggest that the importance of *FRQ* is increasing in the

TABLE 4

The Relation between Financial Reporting Quality and Investment Efficiency

$$InvEff_{i,t+1} = \beta_0 + \beta_1 FRQ_{i,t} + \beta_n Control\ Variables_{i,t} + \sum Industries + \sum Countries + \varepsilon_{i,t}$$

		Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
Variables	Predictions	*FRQ* Measured by *DisAccr* (1)	*FRQ* Measured by *DisRev* (2)	*FRQ* Measured by *DD* (3)	*FRQ* Measured by *Aggreg* (4)	*FRQ* Measured by *DisAccr* (5)	*FRQ* Measured by *DisRev* (6)	*FRQ* Measured by *DD* (7)	*FRQ* Measured by *Aggreg* (8)
FRQ	− (H1)	−0.020***	−0.055***	−0.020***	−0.004***	−0.102***	−0.105***	−0.095***	−0.013***
		(−5.95)	(−6.69)	(−5.84)	(−8.50)	(−4.15)	(−2.73)	(−3.83)	(−4.12)
LogAssets		−0.001**	−0.001*	−0.001*	−0.001**	−0.022***	−0.022***	−0.022***	−0.021***
		(−2.19)	(−1.92)	(−1.74)	(−2.56)	(−6.44)	(−6.32)	(−6.18)	(−6.10)
LogAge		−0.001	0.001	−0.001	0.001	−0.007	−0.007	−0.010	−0.008
		(−0.22)	(0.60)	(−0.37)	(0.06)	(−0.84)	(−0.78)	(−1.17)	(−0.91)
Tang		0.002*	0.002	0.002	0.003**	0.028**	0.025*	0.027**	0.031**
		(1.75)	(1.37)	(1.60)	(2.22)	(2.18)	(1.93)	(2.14)	(2.42)
Slack		0.004	0.001	0.002	0.003	0.011	0.015	0.005	0.005
		(1.24)	(0.26)	(0.56)	(0.81)	(0.44)	(0.59)	(0.21)	(0.21)
Audit		−0.000	−0.001	0.001	0.001	−0.011	−0.013*	−0.010	−0.009
		(−0.08)	(−1.52)	(0.23)	(0.35)	(−1.59)	(−1.91)	(−1.52)	(−1.32)
Intercept		0.046***	0.046***	0.046***	0.048***	0.259***	0.264***	0.261***	0.261***
		(15.91)	(15.63)	(16.08)	(17.62)	(10.87)	(10.92)	(10.90)	(10.89)
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.434	0.463	0.448	0.450	0.143	0.138	0.139	0.140
n		4,653	4,917	4,594	4,533	1,819	1,810	1,794	1,788

---

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, using two-tailed tests.

The table reports OLS regression results. Variables are defined in Appendix A. Huber-White robust standard errors are used to control for heteroscedasticity.

TABLE 5

The Conditional Effect of Bank Financing on the Relation between Financial Reporting Quality and Investment Efficiency

$$InvEffi_{i,t+1} = \beta_0 + \beta_1 FRQ_{i,t} + \beta_2 Bank_{i,t} + \beta_3 FRQ \times Bank_{i,t} + \beta_n Control\ Variables_{i,t} + \sum Industries + \sum Countries + \varepsilon_{i,t}$$

		Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
Variables	**Predictions**	***FRQ* Measured by *DisAccr* (1)**	***FRQ* Measured by *DisRev* (2)**	***FRQ* Measured by *DD* (3)**	***FRQ* Measured by *Aggreg* (4)**	***FRQ* Measured by *DisAccr* (5)**	***FRQ* Measured by *DisRev* (6)**	***FRQ* Measured by *DD* (7)**	***FRQ* measured by *Aggreg* (8)**
FRQ	− (H1)	−0.011**	−0.031**	−0.015***	−0.003***	−0.117***	−0.129**	−0.090***	−0.014***
		(−2.11)	(−2.50)	(−2.94)	(−4.20)	(−3.67)	(−2.49)	(−2.65)	(−3.44)
Bank		0.001	−0.002	0.001	0.002	0.006	0.005	−0.002	−0.002
		(1.20)	(−1.44)	(0.81)	(1.52)	(0.66)	(0.57)	(−0.23)	(−0.22)
FRQ × *Bank*	− (H2)	−0.015**	−0.044**	−0.017**	−0.002**	−0.065	−0.092	−0.001	−0.006
		(−2.06)	(−2.29)	(−2.21)	(−1.98)	(−1.22)	(−1.14)	(−0.03)	(−0.83)
LogAssets		−0.001	−0.001	−0.001	−0.001	−0.023***	−0.023***	−0.022***	−0.022***
		(−1.03)	(−0.78)	(−0.80)	(−1.03)	(−6.00)	(−5.81)	(−5.56)	(−5.57)
LogAge		−0.001	0.001	−0.001	−0.001	−0.009	−0.009	−0.011	−0.009
		(−0.38)	(0.42)	(−1.03)	(−0.10)	(−0.86)	(−0.85)	(−1.12)	(−0.86)
Tang		0.004**	0.003	0.003**	0.005**	0.034**	0.033**	0.035**	0.040***
		(2.28)	(1.58)	(1.96)	(2.57)	(2.35)	(2.26)	(2.42)	(2.71)
Slack		0.010***	0.005	0.009**	0.008**	0.020	0.023	0.022	0.020
		(2.61)	(1.12)	(2.17)	(2.01)	(0.69)	(0.78)	(0.74)	(0.67)
Audit		0.001	−0.002	0.001	0.001	−0.013	−0.016**	−0.013	−0.012
		(0.50)	(−1.56)	(1.02)	(0.96)	(−1.62)	(−1.99)	(−1.63)	(−1.47)
Intercept		0.042***	0.045***	0.042***	0.042***	0.265***	0.270***	0.263***	0.265***
		(11.49)	(11.89)	(12.34)	(12.47)	(10.07)	(10.00)	(9.92)	(10.00)
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.271	0.304	0.275	0.274	0.093	0.088	0.086	0.088
n		2,977	3,200	2,881	2,861	1,397	1,390	1,378	1,373

, * Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, using two-tailed tests.

The table reports OLS regression results. Compared with Table 4, we further require availability of data for a firm's financing sources. Variables are defined in Appendix A. Huber-White robust standard errors are used to control for heteroscedasticity.

presence of bank financing (relative to other external financing sources), which is consistent with banks making use of financial statements in their lending decisions.[21]

In Table 6, we test H3, the effect of tax incentives. Recall that we argue that the informational role of *FRQ* will be reduced when the overriding objective is to minimize taxes. The main effect of *FRQ* continues to be negative and significant in this specification. As predicted, we observe coefficients on the interaction of *FRQ* and *Tax* that are positive across all eight specifications. More specifically, *FRQ* × *Tax* is positive and significant (at the 1 percent level) for all *FRQ* proxies for the overinvestment sample. In other words, consistent with H3, the effect of *FRQ* in mitigating the overinvestment problem is reduced to the extent that firms' incentives are primarily tax-oriented.[22] In the underinvestment scenario we observe positive but statistically weaker coefficients on *FRQ* × *Tax*. Thus, there is at least some support for the often-argued (but seldom-tested) notion in the literature that tax motivations detract from the usefulness of accounting information. This finding adds to the literature insights regarding how financial reporting and tax incentives jointly influence investment decisions; previous studies have generally ignored tax considerations (Hanlon and Heitzman 2010).

Finally, to test whether bank financing and tax incentives have incremental effects in the presence of each other, Table 7 reports tests in which we include both *Bank* and *Tax* jointly (and thus have a smaller sample). As before, the estimated coefficient on *FRQ* × *Bank* continues to be negative in all specifications and statistically significant for the underinvestment scenario. Furthermore, the estimated coefficient on *FRQ* × *Tax* is significantly positive for the overinvestment case but not significant for the underinvestment case. Importantly, the main effect of *FRQ* is yet again significantly negative across all test specifications. In addition to being of interest themselves, our interaction tests also serve the function of providing additional credence to H1 (e.g., related to correlated omitted variables or causality). The estimated coefficients on the two interaction terms have the predicted signs and are statistically significant in several cases. In other words, the effect is more (less) pronounced in subsamples in which we predict the effect to be stronger (weaker). As Rajan and Zingales (1998) point out, it would be difficult to envision a consistent theory in which causality is reversed yet the subsample results hold.[23]

V. ADDITIONAL TESTS

In this section, we report results of additional tests that lend robustness and extend the reported results. For brevity, we only tabulate the results of the alternative investment model specifications.

Alternative Investment Efficiency Model Specifications

We conduct four sensitivity tests related to our measurement of investment efficiency. First, we estimate the investment efficiency model after adding lagged investment to Equation (1). By using

---

[21] Unlike our proxy for tax incentives, which we consider mostly outside of managers' control, the importance of bank financing could be driven by other firm characteristics. For this reason we estimate a two-stage Heckman model in which the first stage (which is based on Beck et al. 2008) predicts *Bank* and the second stage controls for the Inverse Mills Ratio estimated from the first stage. As an instrument we include the approximate value of collateral as a percentage of the loan value (*Collateral*). Banks may be less willing to lend to firms when the amount of assets that can be pledged as collateral is limited. In untabulated tests we find that *Collateral* is significantly correlated with the choice of bank financing but is not significantly associated with over- or underinvestment. After controlling for potential self-selection, our conclusions regarding H2 remain unchanged: the importance of earnings quality for investment efficiency is increasing in bank financing.

[22] In fact, for all four financial reporting quality proxies, the sum of the coefficients on *FRQ* and *FRQ* × *Tax* ($\beta_1 + \beta_3$) is not statistically different from 0 in the overinvestment tests.

[23] For example, the theory would need to explain why, for firms with low bank financing and for firms that are more motivated to manage accounting earnings for tax purposes, relatively higher levels of *FRQ* would result in especially high investment efficiency.

TABLE 6

The Conditional Effect of Tax Burdens on the Relation between Financial Reporting Quality and Investment Efficiency

$$InvEffi_{i,t+1} = \beta_0 + \beta_1 FRQ_{i,t} + \beta_2 Tax_{i,t} + \beta_3 FRQ \times Tax_{i,t} + \beta_n Control\ Variables_{i,t} + \sum Industries + \sum Countries + \varepsilon_{i,t}$$

		Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
Variables	**Predictions**	***FRQ* Measured by *DisAccr* (1)**	***FRQ* Measured by *DisRev* (2)**	***FRQ* Measured by *DD* (3)**	***FRQ* Measured by *Aggreg* (4)**	***FRQ* Measured by *DisAccr* (5)**	***FRQ* Measured by *DisRev* (6)**	***FRQ* Measured by *DD* (7)**	***FRQ* Measured by *Aggreg* (8)**
FRQ	− (H1)	−0.020***	−0.062***	−0.022***	−0.004***	−0.128***	−0.139***	−0.124***	−0.018***
		(−5.89)	(−7.20)	(−5.83)	(−8.47)	(−4.80)	(−3.35)	(−4.34)	(−5.11)
Tax		−0.002	−0.001	−0.001	−0.002*	0.015*	0.011	0.007	−0.005
		(−1.51)	(−0.22)	(−1.02)	(−1.89)	(1.87)	(1.38)	(0.95)	(−0.75)
FRQ × *Tax*	+ (H3)	0.002	0.038*	0.008	0.001	0.148***	0.209***	0.145***	0.023***
		(0.19)	(1.72)	(0.79)	(0.78)	(3.12)	(2.63)	(3.08)	(4.01)
LogAssets		−0.001**	−0.001*	−0.001*	−0.001**	−0.022***	−0.022***	−0.021***	−0.021***
		(−2.03)	(−1.82)	(−1.66)	(−2.47)	(−6.34)	(−6.30)	(−6.01)	(−6.01)
LogAge		−0.001	0.001	−0.001	0.001	−0.007	−0.007	−0.010	−0.008
		(−0.17)	(0.62)	(−0.41)	(0.02)	(−0.84)	(−0.79)	(−1.10)	(−0.88)
Tang		0.002*	0.002	0.002	0.003**	0.026**	0.024*	0.026**	0.031**
		(1.90)	(1.44)	(1.63)	(2.25)	(2.03)	(1.85)	(2.02)	(2.37)
Slack		0.004	0.001	0.002	0.003	0.013	0.017	0.009	0.009
		(1.27)	(0.32)	(0.53)	(0.79)	(0.52)	(0.65)	(0.34)	(0.36)
Audit		0.000	−0.001	0.001	0.001	−0.011	−0.013*	−0.011	−0.010
		(0.02)	(−1.45)	(0.30)	(0.41)	(−1.55)	(−1.85)	(−1.58)	(−1.43)
Intercept		0.046***	0.046***	0.046***	0.048***	0.255***	0.262***	0.255***	0.258***
		(15.71)	(15.45)	(16.02)	(17.59)	(10.65)	(10.79)	(10.62)	(10.78)
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.434	0.464	0.448	0.450	0.146	0.139	0.141	0.144
n		4,628	4,892	4,594	4,533	1,811	1,803	1,786	1,781

---

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, using two-tailed tests.

The table reports OLS regression results. Compared with Table 4, we further require availability of data for a firm's tax burden. Variables are defined in Appendix A. Huber-White robust standard errors are used to control for heteroscedasticity.

TABLE 7

The Conditional Effect of Bank Financing and Tax Burdens on the Relation between Financial Reporting Quality and Investment Efficiency

$$InvEffi_{i,t+1} = \beta_0 + \beta_1 FRQ_{i,t} + \beta_2 Bank_{i,t} + \beta_3 FRQ \times Bank_{i,t} + \beta_4 Tax_{i,t} + \beta_5 FRQ \times Tax_{i,t} + \beta_n Control\ Variables_{i,t} + \sum Industries + \sum Countries + \varepsilon_{i,t}$$

		Dependent Variable = *Underinvestment*				**Dependent Variable = *Overinvestment***			
Variables	**Predictions**	***FRQ* Measured by *DisAccr* (1)**	***FRQ* Measured by *DisRev* (2)**	***FRQ* Measured by *DD* (3)**	***FRQ* Measured by *Aggreg* (4)**	***FRQ* Measured by *DisAccr* (5)**	***FRQ* Measured by *DisRev* (6)**	***FRQ* Measured by *DD* (7)**	***FRQ* Measured by *Aggreg* (8)**
FRQ	− (H1)	−0.011**	−0.037***	−0.014***	−0.003***	−0.158***	−0.176***	−0.116***	−0.019***
		(−2.06)	(−2.85)	(−2.70)	(−4.04)	(−4.56)	(−3.14)	(−3.04)	(−4.15)
Bank		0.002	−0.001	0.001	0.002	0.007	0.006	−0.002	−0.002
		(1.35)	(−1.24)	(0.81)	(1.34)	(0.77)	(0.73)	(−0.26)	(−0.24)
FRQ × *Bank*	− (H2)	−0.014*	−0.042**	−0.017**	−0.002**	−0.079	−0.118	−0.001	−0.007
		(−1.95)	(−2.14)	(−2.22)	(−1.98)	(−1.52)	(−1.47)	(−0.02)	(−1.10)
Tax		−0.002	−0.001	−0.002	−0.002	0.017*	0.009	0.004	−0.007
		(−1.56)	(−0.59)	(−1.51)	(−1.62)	(1.82)	(0.95)	(0.41)	(−0.96)
FRQ × *Tax*	+ (H3)	−0.003	0.020	−0.003	0.000	0.191***	0.228***	0.125**	0.021***
		(−0.23)	(0.71)	(−0.24)	(0.08)	(3.66)	(2.57)	(2.33)	(2.96)
LogAssets		−0.001	−0.001	−0.001	−0.001	−0.023***	−0.023***	−0.021***	−0.021***
		(−1.00)	(−0.78)	(−0.79)	(−1.02)	(−5.85)	(−5.75)	(−5.31)	(−5.38)
LogAge		−0.001	0.001	−0.001	−0.000	−0.009	−0.009	−0.011	−0.009
		(−0.28)	(0.49)	(−1.03)	(−0.11)	(−0.87)	(−0.86)	(−1.08)	(−0.87)
Tang		0.005**	0.003*	0.003**	0.005**	0.032**	0.033**	0.034**	0.039***
		(2.42)	(1.69)	(1.96)	(2.57)	(2.22)	(2.27)	(2.32)	(2.68)
Slack		0.011***	0.005	0.008**	0.008**	0.022	0.025	0.024	0.022
		(2.67)	(1.19)	(2.15)	(1.99)	(0.76)	(0.86)	(0.81)	(0.74)
Audit		0.001	−0.002	0.001	0.001	−0.013	−0.016**	−0.014*	−0.013
		(0.55)	(−1.54)	(1.04)	(0.98)	(−1.62)	(−2.01)	(−1.70)	(−1.60)

(continued on next page)

TABLE 7 (continued)

		Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
Variables	Predictions	***FRQ*** Measured by *DisAccr* (1)	***FRQ*** Measured by *DisRev* (2)	***FRQ*** Measured by *DD* (3)	***FRQ*** Measured by *Aggreg* (4)	***FRQ*** Measured by *DisAccr* (5)	***FRQ*** Measured by *DisRev* (6)	***FRQ*** Measured by *DD* (7)	***FRQ*** Measured by *Aggreg* (8)
Intercept		0.041*** (11.40)	0.044*** (11.79)	0.042*** (12.39)	0.042*** (12.51)	0.259*** (9.81)	0.266*** (9.80)	0.257*** (9.63)	0.262*** (9.87)
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.270	0.303	0.276	0.274	0.098	0.089	0.088	0.092
n		2,955	3,178	2,881	2,861	1,390	1,384	1,371	1,367

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, using two-tailed tests.

The table reports OLS regression results. Compared with Table 4, we further require availability of data for firms' financing sources, and tax burden. Variables are defined in Appendix A. Huber-White robust standard errors are used to control for heteroscedasticity.

such a "change specification," we in effect use the firm as its own control, mitigating the possibility of omitted variables from the model. However, this control comes at the cost of a reduced sample size.

As our second test, we measure investment efficiency using an expanded model based on Richardson (2006). We estimate optimal investment according to the following regression specification:

$$\begin{aligned} Invest_{i,t} = {} & \alpha_0 + \alpha_1 NEG_{i,t-1} + \alpha_2 \%RevGrowth_{i,t-1} + \alpha_3 NEG * \%RevGrowth_{i,t-1} \\ & + \alpha_4 LogAssets_{i,t-1} + \alpha_5 LogAge_{i,t-1} + \alpha_6 Slack_{i,t-1} + \alpha_7 Lev_{i,t-1} \\ & + \alpha_8 BConIndex_{i,t} + \alpha_9 LConIndex_{i,t} + \alpha_{10} PEConIndex_{i,t} + \varepsilon_{i,t}. \end{aligned} \qquad (6)$$

In addition to the proxy for growth opportunity, we also include firm size, firm age, financial slack, and financial leverage as controls for financing constraints. Moreover, based on Brown et al. (2008), we augment the model with three additional firm-level variables related to constraints in the operating environment: business constraints, legal constraints, and political/economic constraints.[24]

Third, we replace revenue growth with asset growth (*AssetGrowth*) as our proxy for investment opportunities. Specifically, we follow McNichols and Stubben (2008, 1579) and use the natural log of total assets at the end of year $t-1$ divided by total assets at the end of year $t-2$ for the estimation of Equation (1).

Finally, we follow Biddle et al. (2009) and sort firms, based on the positive and negative residuals, respectively, into deciles. Within the positive-residual group, we remove the bottom decile because these firms, whose unexpected investments are closest to 0 among all overinvesting firms, are more likely to be affected by measurement error in the investment model (i.e., misclassified as overinvesting firms). Similarly, we remove the top decile from the negative-residual group. We then repeat all the tests using the remaining observations.

The results of these four tests are tabulated in Panels A through D of Table 8. Even with the smaller sample sizes in these tests, results are similar to those of the main test specifications, and no inferences are affected.

Additional Control Variables

Although we include several firm-level control variables as well as industry and country fixed effects in our main tests, it is always possible that there are some omitted (and correlated) variables. However, adding more control variables comes at the cost of reducing the sample size and, thus, a trade-off exists between sample size (and hence generalizability) and "model completeness." In untabulated tests, we add several additional controls to all regressions.

Following Biddle et al. (2009), we include (1) long-term liabilities divided by total assets (*Lev*); (2) the standard deviation of sales in the past three years (*StdRev*); (3) the standard deviation of investments in the past three years (*StdInvest*); (4) the operating cycle calculated by using both accounts receivables and sales (*OperCycle*); and (5) an indicator variable referring to whether a firm reports a loss (*Loss*). We also (6) control for profitability (*ROA*). Furthermore, we take advantage of the availability of certain unique data items from WBES and include: (7) an indicator variable referring to whether the firm exports (*Export*); (8) an indicator variable for whether any foreign company or individual has ownership interests in the firm (*Foreign*); (9) an indicator variable that

---

[24] Appendix A contains detailed definitions of these additional variables. Note that with the expanded set of control variables, we require more observations to estimate the model than for the estimation of Equation (1). Therefore, we estimate this model by country and adjust both dependent and independent variables by the country-specific industry median. The adjustment for industry median accounts for industry heterogeneity.

TABLE 8

Sensitivity Analyses Related to the Investment Efficiency Model

$$InvEff_{i,t+1} = \beta_0 + \beta_1 FRQ_{i,t} + \beta_2 Bank_{i,t} + \beta_3 FRQ \times Bank_{i,t} + \beta_4 Tax_{i,t} + \beta_5 FRQ \times Tax_{i,t} + \beta_n Control\ Variables_{i,t} + \sum Industries + \sum Countries + \varepsilon_{i,t}$$

Panel A: Alternative Investment Efficiency Model with Lagged Investment

		Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
Variables	**Predictions**	***FRQ* Measured by *DisAccr* (1)**	***FRQ* Measured by *DisRev* (2)**	***FRQ* Measured by *DD* (3)**	***FRQ* Measured by *Aggreg* (4)**	***FRQ* Measured by *DisAccr* (5)**	***FRQ* Measured by *DisRev* (6)**	***FRQ* Measured by *DD* (7)**	***FRQ* Measured by *Aggreg* (8)**
FRQ	− (H1)	−0.016***	−0.022*	−0.0167***	−0.003***	−0.080***	−0.200***	−0.059*	−0.013***
		(−2.62)	(−1.68)	(−2.69)	(−3.26)	(−2.82)	(−4.11)	(−1.93)	(−3.49)
Bank		−0.001	−0.003***	−0.001	0.001	0.002	0.008	−0.002	−0.002
		(−0.22)	(−2.59)	(−0.39)	(1.02)	(0.30)	(1.23)	(−0.31)	(−0.34)
FRQ × *Bank*	− (H2)	−0.015*	−0.056***	−0.018**	−0.003**	−0.004	−0.111	−0.003	−0.003
		(−1.77)	(−2.72)	(−2.01)	(−2.46)	(−0.10)	(−1.45)	(−0.08)	(−0.56)
Tax		−0.001	0.001	−0.002	−0.001	0.004	0.008	−0.002	−0.008
		(−0.77)	(0.35)	(−1.55)	(−0.49)	(0.56)	(1.05)	(−0.24)	(−1.25)
FRQ × *Tax*	+ (H3)	−0.006	0.015	0.022	0.001	0.099**	0.254***	0.073*	0.016***
		(−0.38)	(0.58)	(1.35)	(0.54)	(2.20)	(3.12)	(1.68)	(2.67)
Control variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.196	0.238	0.198	0.198	0.103	0.107	0.098	0.103
n		2,836	3,042	2,788	2,768	1,337	1,332	1,301	1,296

(continued on next page)

TABLE 8 (continued)

Panel B: Alternative Investment Efficiency Model with Additional Control Variables

		Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
Variables	**Predictions**	***FRQ*** **Measured by** ***DisAccr*** **(1)**	***FRQ*** **Measured by** ***DisRev*** **(2)**	***FRQ*** **Measured by** ***DD*** **(3)**	***FRQ*** **Measured by** ***Aggreg*** **(4)**	***FRQ*** **Measured by** ***DisAccr*** **(5)**	***FRQ*** **Measured by** ***DisRev*** **(6)**	***FRQ*** **Measured by** ***DD*** **(7)**	***FRQ*** **Measured by** ***Aggreg*** **(8)**
FRQ	− (H1)	−0.019***	−0.011*	−0.018**	−0.003***	−0.114***	−0.157**	−0.104***	−0.016***
		(−2.74)	(−1.66)	(−2.39)	(−2.71)	(−3.14)	(−2.46)	(−2.62)	(−3.17)
Bank		0.002	−0.001	0.001	0.003	0.009	0.008	−0.001	−0.002
		(1.26)	(−0.02)	(0.81)	(1.50)	(0.98)	(0.73)	(−0.13)	(−0.34)
FRQ × *Bank*	− (H2)	−0.011*	−0.048**	−0.017*	−0.002**	−0.090	−0.096	−0.021	−0.008
		(−1.72)	(−2.04)	(−1.78)	(−2.05)	(−1.43)	(−1.02)	(−0.35)	(−1.05)
Tax		0.003	0.003	0.002	0.003**	0.009	0.002	0.001	−0.008
		(1.56)	(1.62)	(0.93)	(2.00)	(0.93)	(0.20)	(0.04)	(−1.06)
FRQ × *Tax*	+ (H3)	−0.001	0.002	0.014	−0.001	0.141***	0.185**	0.095*	0.016*
		(−0.02)	(0.05)	(0.92)	(−0.28)	(2.67)	(2.02)	(1.73)	(1.95)
Control variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.239	0.251	0.248	0.247	0.083	0.149	0.076	0.079
n		2,177	2,308	2,128	2,114	1,167	1,168	1,138	1,135

(continued on next page)

TABLE 8 (continued)

Panel C: Alternative Investment Efficiency Model with the Asset Growth Variable

		Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
Variables	**Predictions**	***FRQ* Measured by *DisAccr* (1)**	***FRQ* Measured by *DisRev* (2)**	***FRQ* Measured by *DD* (3)**	***FRQ* Measured by *Aggreg* (4)**	***FRQ* Measured by *DisAccr* (5)**	***FRQ* Measured by *DisRev* (6)**	***FRQ* Measured by *DD* (7)**	***FRQ* Measured by *Aggreg* (8)**
FRQ	− (H1)	−0.028***	−0.011*	−0.031***	−0.004***	−0.135***	−0.194***	−0.101***	−0.017***
		(−4.50)	(−1.77)	(−4.68)	(−4.85)	(−4.02)	(−3.44)	(−2.76)	(−3.76)
Bank		0.001	−0.002	−0.001	0.001	0.002	0.005	−0.005	−0.006
		(0.18)	(−1.55)	(−0.54)	(1.17)	(0.21)	(0.51)	(−0.56)	(−0.84)
FRQ × *Bank*	− (H2)	−0.013*	−0.032**	−0.018**	−0.002*	−0.058	−0.125	−0.011	−0.006
		(−1.65)	(−2.00)	(−2.08)	(−1.86)	(−1.10)	(−1.57)	(−0.20)	(−0.93)
Tax		−0.003**	−0.003*	−0.004**	−0.001	0.019**	0.009	0.007	−0.005
		(−2.09)	(−1.84)	(−2.45)	(−1.02)	(2.02)	(0.99)	(0.82)	(−0.60)
FRQ × *Tax*	+ (H3)	0.018	0.041	0.019	0.003	0.193***	0.195**	0.131**	0.018**
		(1.39)	(1.27)	(1.16)	(1.51)	(3.76)	(2.16)	(2.46)	(2.54)
Control variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.274	0.291	0.273	0.234	0.092	0.086	0.082	0.085
n		2,948	3,171	2,893	2,873	1,397	1,391	1,359	1,355

(continued on next page)

TABLE 8 (continued)

Panel D: Alternative Investment Efficiency Measure after Removing the Deciles around the Zero Residuals

Variables	Predictions	Dependent Variable = *Underinvestment*				Dependent Variable = *Overinvestment*			
		FRQ Measured by *DisAccr* (1)	*FRQ* Measured by *DisRev* (2)	*FRQ* Measured by *DD* (3)	*FRQ* Measured by *Aggreg* (4)	*FRQ* Measured by *DisAccr* (5)	*FRQ* Measured by *DisRev* (6)	*FRQ* Measured by *DD* (7)	*FRQ* Measured by *Aggreg* (8)
FRQ	− (H1)	−0.015***	−0.045***	−0.016***	−0.003***	−0.158***	−0.187***	−0.114***	−0.019***
		(−2.91)	(−3.64)	(−3.15)	(−4.70)	(−4.56)	(−3.33)	(−3.02)	(−4.12)
Bank		0.002	−0.002	0.002	0.002	0.007	0.009	0.001	−0.001
		(1.23)	(−1.50)	(1.34)	(1.57)	(0.77)	(1.01)	(0.05)	(−0.07)
FRQ × *Bank*	− (H2)	−0.010*	−0.047***	−0.016**	−0.002*	−0.079	−0.116	−0.027	−0.011
		(−1.70)	(−2.70)	(−2.08)	(−1.90)	(−1.52)	(−1.42)	(−0.47)	(−1.50)
Tax		−0.001	0.000	−0.001	−0.001	0.017*	0.006	0.000	−0.008
		(−0.70)	(0.12)	(−0.99)	(−0.99)	(1.82)	(0.58)	(0.02)	(−1.05)
FRQ × *Tax*	+ (H3)	−0.001	0.019	−0.003	0.001	0.191***	0.187**	0.093*	0.018**
		(−0.09)	(0.81)	(−0.27)	(0.01)	(3.66)	(2.07)	(1.75)	(2.41)
Control variables		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2		0.292	0.331	0.291	0.292	0.083	0.076	0.073	0.077
n		2,754	2,977	2,725	2,705	1,390	1,283	1,270	1,266

*, **, *** Represent significance at the 10 percent, 5 percent, and 1 percent levels, respectively, using two-tailed tests.

The table reports OLS regression results that replicate results in Table 7 with four alternative investment efficiency models. Variables are defined in Appendix A. In Panel A, we compute investment efficiency by adding the lagged investment variable to Equation (1). In Panel B, we adopt an expanded investment efficiency model from Equation (4). The expanded model includes additional control variables, such as firm size, firm age, financial slack, financial leverage, business constraints, legal constraints, and political/economic constraints. In Panel C, we compute investment efficiency by replacing revenue growth with the asset growth variable in Equation (1). In Panel D, we compute investment efficiency by removing decile observations surrounding the zero residuals of Equation (1). Huber-White robust standard errors are used to control for heteroscedasticity.

takes on the value of 1 if financing obstacles (*FinCon*) are severe; and (10) an indicator variable that takes on the value of 1 if the largest shareholder owns more than 50 percent of the firm (*Control*).

Requiring data on these additional controls reduces the sample size by approximately half. Nevertheless, even with the reduced sample size our conclusions for H1 to H3 remain unaltered.

Controls for Potential Endogeneity

Our tests are based on the premise that FRQ affects investment efficiency. We acknowledge that it is challenging to establish causality in this line of research. However, we have implemented research design features to at least partially alleviate such concerns. First, and perhaps most important, economic theory supports our finding that FRQ has a positive effect on investment decisions. We are unaware of theory suggesting a reverse relation. Second, we test the effect of FRQ in the current period on *next* period's investment efficiency. Third, we include control variables motivated by prior research. In addition, above we introduce a number of additional controls. Fourth, our focus on interaction effects makes it hard to argue for reverse causality (e.g., Rajan and Zingales 1998).

As a final control for endogeneity of *FRQ*, we consider a two-stage least squares model. To model cross-sectional variation in *FRQ*, we first include all control variables from Equation (5). Then we add a variable that reflects the strength of the relation between the firm and local suppliers and customers (or the firm's reliance on local product markets). Extant research (e.g., Bowen et al. 1995; Raman and Shahrur 2008; Dou et al. 2011) shows that a firm's suppliers and customers affect the firm's *FRQ*.[25] The instrument (*LocalMarket*) is constructed by averaging the percentage of domestic sales of outputs and the percentage of domestic purchases of inputs. *LocalMarket* is significantly correlated with all four measures of *FRQ* but is not significantly correlated with investment efficiency.[26] Untabulated results show that no inferences are affected after controlling for potential endogeneity of *FRQ* through this two-stage estimation.

Alternative Country Controls and Possible Variations across Sample Countries

Although all sample countries are emerging economies, they clearly differ from each other in various ways. Recall that we control for such effects using country fixed effects in all reported tests (e.g., Doidge et al. 2007). In this section, we perform two additional tests. First, we replace country fixed effects with legal origin (i.e., English common law versus other legal traditions) and creditor protection, either separately or jointly. These variables come from Djankov et al. (2007) and have been widely used in recent accounting and finance research. No conclusions are altered with this specification. Second, we separate the sample based on legal origins and the values of creditor rights, respectively, and estimate the regressions separately for the two groups.[27] For both partitions, we find that our results are qualitatively unchanged for both sets of countries.

We have also estimated pooled regressions and included both the main of legal origin and creditor rights and their interactions with FRQ. It is interesting to note that the interaction terms are always negative, although only statistically significant for the overinvestment scenario. These results are consistent with prior research that suggests that FRQ is more important in countries with stronger investor protection.

---

[25] In addition, this instrument reflects the monitoring costs of local stakeholders relative to the costs by foreign stakeholders (e.g., Lerner 1995).

[26] The first-stage model has adjusted R^2s of 21 percent, 13 percent, 15 percent, and 22 percent for the four models, respectively.

[27] As the creditor rights index ranges from 0 (weak) to 4 (strong), we classify any value 0–1 into the low creditor protection group, and any value 2–4 into the high creditor protection group.

Sensitivity Analyses Related to the Degree of Accrual Accounting Employed

By construction, our sample firms use accrual accounting (i.e., we require data on balance sheet items necessary to compute accrual-based FRQ measures). However, to address the possibility that some firms use accrual accounting to a lesser extent, we conduct the following tests. First, as the smallest private firms are likely to rely less on accrual accounting than larger firms (Allee and Yohn 2009), we repeat the tests after excluding the smallest firms. We run the following four tests: (1) exclude the smallest (based on total assets) 10 percent of firms for the pooled sample; (2) exclude the smallest 20 percent of firms for the pooled sample; (3) exclude the smallest 10 percent of firms by country; and (4) exclude the smallest 20 percent of firms by country. In these samples, which are more likely to generalize to other private firms that also use accrual accounting, our results are qualitatively unchanged. Second, insofar as Allee and Yohn (2009) find that most U.S. private firms that are audited use accrual accounting, we repeat the tests keeping only firms that are audited. Again, inferences are unaffected.

Finally, country-level regulations of private firms' financial reporting differ in their strength across our sample countries. We manually collect relevant information and construct a country-level indicator of financial reporting regulation of private firms, *FinRegulation*, set equal to 1 if the sample firm is from a country that imposes substantial financial reporting regulation on private firms, and 0 otherwise.[28] We first add *FinRegulation* as an additional control in our main test. The coefficients on the FRQ proxies still load significantly and with the expected negative signs.[29]

The Importance of Financial Reporting Quality for Public Firms Included in the WBES Database

Although the focus in our study is on private firms from emerging markets, the WBES database also contains a relatively small number of public firms. Only 453 (130) underinvestment (overinvestment) observations remain when using public firms and requiring data availability as described above. In this small sample, we find that, consistent with prior research using our private firm results, FRQ is negatively related to both under- and overinvestment.[30]

Additional Tests Related to Cash Constraints

Finally, motivated by prior research (e.g., Jensen 1986; Biddle et al. 2009), we examine the role of FRQ for firms that are above and those that are below the median for cash constraints (defined as the cash balance divided by total assets) separately. Our results suggest that FRQ is positively associated with investment efficiency for both sets of firms.

VI. CONCLUSION

We study the role of accounting information for a set of firms for which there is very limited prior research evidence: private firms from emerging markets. Although private firms make up the vast majority of economic activity around the world, they have received limited attention in academic research, and in particular we are unaware of prior research on the relation between financial reporting quality and investment efficiency in a private firm setting.

---

[28] Our classification scheme suggests that Bangladesh, Ecuador, Eritrea, Guatemala, Honduras, Nicaragua, Oman, Philippines, Sri Lanka, and Syria appear to have no or very limited financial reporting regulation of private firms.

[29] In additional tests, we also include an interaction term between *FRQ* and *FinRegulation*. This interaction term is significantly negative for underinvestment but not significant for overinvestment. For all specifications, *FRQ* retains its statistical significance.

[30] Specifically, we find that the estimated coefficients on *FRQ* are statistically significant in all specifications when country fixed effects are excluded and are negative but not always significant when country fixed effects are included. This result is not surprising given the limited within-country variation in this small sample.

Prior research suggests that private firms have lower quality financial reporting, presumably due to reduced market demand. In addition, research indicates that firms from countries with less sophisticated institutions also have lower quality accounting. Thus, our study examines the "boundary condition" (or lower bound) in which FRQ may not have the same effect on investment efficiency as that documented for publicly traded companies in the United States.

Research shows that small- and medium-sized enterprises (SMEs) incur not only higher financing obstacles than large firms, but also that the effect of these financing constraints is stronger for SMEs than for large firms (Beck and Demirguc-Kunt 2006). Policymakers in governmental and international aid organizations believe that, in developing countries, small firms have inadequate access to external financing due to market imperfections (World Bank 2007). In other words, compared with public firms in developed countries, our sample firms are more likely to be underinvestment firms, and our empirical evidence supports this idea. Given the importance of financing for these firms, notwithstanding the arguments put forth in prior research, FRQ could play an important role in their investment decisions, and our results are consistent with this notion.

We find evidence that FRQ is positively associated with investment efficiency for our sample firms. Although they may well have lower FRQ than public firms in developed markets, there are more limited alternative information sources other than accounting for private firms (and in countries with lower disclosure levels, on average). Also, the link between financial accounting and management accounting is likely to be strong in these settings, suggesting that the financial accounting information set is largely the same set used in managers' decisions. Further, compared with publicly traded companies in the U.S. and other developed countries, our sample firms face fewer mandatory accounting recognition and measurement (as well as disclosure) requirements. Therefore, our sample firms have greater flexibility to choose FRQ and their choices have the potential to be especially important for improving investment efficiency.

In addition, we hypothesize that the source of financing as well as firms' tax incentives affect the role of accounting information. We find evidence supporting these hypotheses. Specifically, we find that greater use of bank financing increases the role that accounting information plays. In addition, for firms in which tax incentives are likely to dominate incentives to provide useful information for internal decision making as well as a source of information for outside providers of capital, the informational role of accounting is significantly diminished. Such a connection between tax-minimization incentives and the informational role of earnings has often been asserted in the literature, but, to date, there is limited empirical evidence on this issue.

Our study is subject to some caveats. First, we rely on survey data from the World Bank. Although prior research provides several validity tests of these data, survey data are known to have limitations. Second, we acknowledge that both the investment efficiency and the FRQ variables likely suffer from measurement error. Third, our results do not necessarily generalize to all private firms. Fourth, it is difficult to prove causality.[31] Notwithstanding these potential limitations, we hope that our findings will be of interest to researchers, standard-setters and other regulators, government officials in emerging markets, managers, and, importantly, the World Bank and others involved in improving economic conditions in developing countries.

---

[31] However, our empirical results are supported by economic theory. Further, we employ research design features that increase our confidence in the inferences drawn. We report results using various specifications of both dependent and test variables, and our tests include numerous control variables motivated by prior research. We further report results of changes tests in addition to association tests, and also report results using 2SLS. Finally, the interaction effects help alleviate concerns over potential omitted variables.

REFERENCES

Abu-Nassar, M., and B. A. Rutherford. 1996. External users of financial reports in less developed countries: The case of Jordan. *British Accounting Review* 28 (1): 73–87.

Alford, A., J. Jones, R. Leftwich, and M. Zmijewski. 1993. The relative informativeness of accounting disclosures in different countries. *Journal of Accounting Research* 31 (Supplement): 183–223.

Ali, A., and L.-S. Hwang. 2000. Country-specific factors related to financial reporting and the value relevance of accounting data. *Journal of Accounting Research* 38 (1): 1–21.

Allee, K. D., and T. L. Yohn. 2009. The demand for financial statements in an unregulated environment: An examination of the production and use of financial statements by privately held small businesses. *The Accounting Review* 84 (1): 1–25.

Al-Razeen, A., and Y. Karbhari. 2007. An empirical investigation into the importance, use, and technicality of Saudi annual corporate information. *Advances in International Accounting* 20 (1): 55–74.

Atwood, T. J., M. S. Drake, and L. A. Myers. 2010. Book-tax conformity, earnings persistence and the association between earnings and future cash flows. *Journal of Accounting and Economics* 50 (1): 111–125.

Ball, R., S. P. Kothari, and A. Robin. 2000. The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics* 29 (1): 1–51.

Ball, R., and L. Shivakumar. 2005. Earnings quality in UK private firms: Comparative loss recognition timeliness. *Journal of Accounting and Economics* 39 (1): 83–128.

Beck, T., and A. Demirguc-Kunt. 2006. Small and medium-size enterprises: Access to finance as a growth constant. *Journal of Banking and Finance* 30 (11): 2931–2943.

Beck, T., A. Demirguc-Kunt, and V. Maksimovic. 2005. Financial and legal constraints to firm growth: Does size matter? *Journal of Finance* 60 (1): 137–177.

Beck, T., A. Demirguc-Kunt, and V. Maksimovic. 2006. The influence of financial and legal institutions on firm size. *Journal of Banking and Finance* 30 (11): 2995–3015.

Beck, T., A. Demirguc-Kunt, and V. Maksimovic. 2008. Financing patterns around the world: Are small firms different? *Journal of Financial Economics* 89 (3): 467–487.

Berry, A., S. Faulkner, M. Hughes, and R. Jarvis. 1993. Financial information, the banker and the small business. *British Accounting Review* 25 (2): 131–150.

Berry, A., P. Grant, and R. Jarvis. 2004. European bank lending to the UK SME sector. *International Small Business Journal* 22 (2): 115–130.

Berzins, J., Ø. Bøhren, and P. Rydland. 2008. Corporate finance and governance in firms with limited liability: Basic characteristics. Working paper, Norwegian School of Management.

Biddle, G., and G. Hilary. 2006. Accounting quality and firm-level capital investment. *The Accounting Review* 81 (5): 963–982.

Biddle, G., G. Hilary, and R. S. Verdi. 2009. How does financial reporting quality relate to investments efficiency? *Journal of Accounting and Economics* 48 (2–3): 112–131.

Blanchard, O., F. Lopez-de-Silanez, and A. Shleifer. 1994. What do firms do with cash windfalls? *Journal of Financial Economics* 36 (3): 337–360.

Bova, F., Y. Dou, and O.-K. Hope. 2011. Employee ownership and firm disclosure. Working paper, University of Toronto.

Bowen, R. M., L. DuCharme, and D. Shores. 1995. Stakeholders' implicit claims and accounting method choice. *Journal of Accounting and Economics* 20 (3): 255–295.

Brown, G. W., L. W. Chavis, and L. F. Klapper. 2008. A new lease on life: Institutions, external financing, and business growth. Working paper, The University of North Carolina at Chapel Hill and World Bank.

Burgstahler, D., L. Hail, and C. Leuz. 2006. The importance of reporting incentives: Earnings management in European private and public firms. *The Accounting Review* 81 (5): 983–1016.

Bushman, R., and A. Smith. 2001. Financial accounting information and corporate governance. *Journal of Accounting and Economics* 31 (1–3): 237–333.

Carlin, W., M. E. Schaffer, and P. Seabright. 2007. Where are the real bottlenecks? Evidence from 20,000 firms in 60 countries about the shadow costs of constraints to firm performance. Working paper, CEPR.

Cummins, J. G., K. A. Hassett, and R. O. Hubbard. 1994. A reconsideration of investment behavior using tax reforms as natural experiments. *Brookings Papers on Economic Activity* 25 (2): 1–74.

Danos, P., D. Holt, and E. A. Imhoff. 1989. The use of accounting information in bank lending decisions. *Accounting, Organizations and Society* 14 (3): 235–246.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 35–59.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *Journal of Finance* 46 (4): 1325–1359.

Djankov, S., C. Mcliesh, and A. Shleifer. 2007. Private credit in 129 countries. *Journal of Financial Economics* 84 (2): 299–329.

Doidge, C., A. Karolyi, and R. Stulz. 2007. Why do countries matter so much for corporate governance? *Journal of Financial Economics* 86 (1): 1–39.

Dou, Y., O.-K. Hope, and W. B. Thomas. 2011. Relationship-specificity, contract enforceability, and income smoothing. Working paper, University of Toronto and The University of Oklahoma.

Drury, C., and M. Tayles. 1995. Issues arising from surveys of management accounting practice. *Management Accounting Research* 6 (3): 267–280.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* 59 (4): 1553–1583.

Eberly, J. C. 1997. International evidence on investment and fundamentals. *European Economic Review* 41 (6): 1055–1078.

Fama, E. F., and M. C. Jensen. 1983. Separation of ownership and controls. *Journal of Law and Economics* 26 (2): 301–325.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen. 1988. Investment and finance reconsidered. *The Brookings Papers on Economic Activity* 19 (1): 141–195.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39 (2): 295–327.

Givoly, D., C. Hayn, and S. P. Katz. 2010. Does public ownership of equity improve earnings quality? *The Accounting Review* 85 (1): 195–225.

Hail, L., and C. Leuz. 2006. International differences in the cost of equity capital: Do legal institutions and securities regulation matter? *Journal of Accounting Research* 44 (3): 485–531.

Hanlon, M., and S. Heitzman. 2010. A review of tax research. *Journal of Accounting and Economics* (December): 127–178.

Hope, O.-K. 2003. Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: An international study. *Journal of Accounting Research* 41 (2): 235–272.

Hope, O.-K., and W. B. Thomas. 2008. Managerial empire building and firm disclosure. *Journal of Accounting Research* 46 (3): 591–626.

Hope, O.-K., W. B. Thomas, and D. Vyas. 2011. Financial credibility, ownership, and financing constraints in private firms. *Journal of International Business Studies* (forthcoming).

Hubbard, R. G. 1998. Capital-market imperfections and investment. *Journal of Economic Literature* 36 (1): 193–225.

Hung, M. 2000. Accounting standards and value relevance of earnings: An international analysis. *Journal of Accounting and Economics* 30 (3): 401–420.

Indjejikian, R., and M. Matejka. 2009. CFO fiduciary responsibilities and annual bonus incentives. *Journal of Accounting Research* 47 (4): 1061–1093.

International Bank for Reconstruction and Development (IBRD). 2003. *Productivity and Investment Climate Survey (PICS): Implementation Manual.* Washington, D.C.: The World Bank Group.

Jensen, M. 1986. Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review* 76 (2): 323–329.

Kitindi, E. G., B. S. Magembe, and A. Sethibe. 2007. Lending decision making and financial information: The usefulness of corporate annual reports to lenders in Botswana. *The International Journal of Applied Economics and Finance* 1 (2): 55–66.

Kothari, S. P., A. J. Leone, and C. E. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.

Lambert, R., C. Leuz, and R. E. Verrecchia. 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45 (2): 385–420.

Lerner, J. 1995. Venture capitalists and the oversight of private firms. *Journal of Finance* 50 (1): 301–318.

Leuz, C., D. Nanda, and P. Wysocki. 2003. Earnings management and investor protection: An international comparison. *Journal of Financial Economics* 69 (3): 505–527.

Leuz, C., and R. E. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38 (Supplement): 91–124.

McNichols, M. F. 2002. Discussion of the quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 61–69.

McNichols, M. F., and S. R. Stubben. 2008. Does earnings management affect firms' investment decisions? *The Accounting Review* 83 (6): 1571–1603.

Mirshekary, S., and S. M. Saudagaran. 2005. Perceptions and characteristics of financial statement users in developing countries: Evidence from Iran. *Journal of International Accounting, Auditing and Taxation* 14 (1): 33–54.

Modigliani, F., and M. Miller. 1958. The cost of capital, corporation finance, and the theory of investment. *American Economic Review* 48 (3): 261–297.

Morck, R., A. Shleifer, and R. W. Vishny. 1988. Management ownership and market valuation: An empirical analysis. *Journal of Financial Economics* 20 (1–2): 293–315.

Myers, S. C., and N. S. Majluf. 1984. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics* 13 (2): 187–221.

Nagar, V., K. Petroni, and D. Wolfenzon. 2011. Governance problems in closely-held corporations. *Journal of Financial and Quantitative Analysis* (forthcoming).

Pacter, P. 2004. Financial reporting: SMEs—Will the GAAP widen for SMEs? *Accountancy* 133 (1325): 118–119.

Rajan, R. G., and L. Zingales. 1998. Financial dependence and growth. *American Economic Review* 88 (3): 559–586.

Raman, K., and H. Shahrur. 2008. Relationship-specific investments and earnings management: Evidence on corporate suppliers and customers. *The Accounting Review* 83 (4): 1041–1081.

Richardson, S. 2006. Over-investment of free cash flow. *Review of Accounting Studies* 11 (2–3): 159–189.

Srinidhi, B. N., and F. A. Gul. 2007. The differential effects of auditors' nonaudit and audit fees on accrual quality. *Contemporary Accounting Research* 24 (2): 595–629.

Stubben, S. 2010. Discretionary revenues as a measure of earnings management. *The Accounting Review* 85 (2): 695–717.

Van Tendeloo, B., and A. Vanstraelen. 2008. Earnings management and audit quality in Europe: Evidence from the private client segment market. *European Accounting Review* 17 (3): 447–469.

World Bank. 2007. *Review of Small Business Activities. World Bank Report.* Washington, D.C.: World Bank.

APPENDIX A

Definitions of Variables

Invest = the sum of new investment in machinery, equipment, vehicles, buildings, land, and R&D activities, less the sale of fixed assets in the current year, scaled by the lagged total assets;

%RevGrowth = the yearly percent growth rate of sales from year $t-2$ to year $t-1$;

TAccr = total accruals, measured as the change in non-cash current assets minus the change in current non-interest-bearing liabilities, minus depreciation and amortization expense, scaled by lagged total assets;

(continued on next page)

APPENDIX A (continued)

TCAccr = total current accruals, measured as the change in non-cash current assets minus the change in current non-interest-bearing liabilities, scaled by lagged total assets;

InvEff = a measure of investment efficiency, which is measured as the absolute values of the residuals from the investment model: $Invest_{i,t} = a_0 + a_1 NEG_{i,t-1} + a_2 \%RevGrowth_{i,t-1} + a_3 NEG * \%RevGrowth_{i,t-1} + \varepsilon_{i,t}$. The indicator variable $NEG_{i,t-1}$ takes the value of 1 for negative revenue growth, and 0 otherwise. The investment model is estimated cross-sectionally with at least eight observations in each industry by country. To mitigate the influence of outliers, we winsorize all variables at the 1 percent and 99 percent levels;

Overinvestment = positive residuals from the investment efficiency model;

Underinvestment = absolute value of the negative residuals from the investment efficiency model;

FRQ (measured by DisAccr) = the absolute residual of the Kothari et al. (2005) discretionary accrual model, multiplied by −1;

FRQ (measured by DisRev) = the absolute residual of the discretionary revenues model as presented in McNichols and Stubben (2008), multiplied by −1;

FRQ (measured by DD) = the absolute residual of the modified Dechow-Dichev model as implemented by Francis et al. (2005) and Srinidhi and Gul (2007), multiplied by −1;

FRQ (measured by Aggreg) = an aggregate financial reporting metric, measured as the average of the standardized previous three measures (i.e., *DisAccr, DisRev*, and *DD*);

Bank = an indicator variable that takes on the value of 1 if bank financing accounts for the highest percentage of total financing among all external sources of financing, and its percentage is larger than 5 percent among all financing sources, and 0 otherwise; and

Tax = an indicator variable that takes on the value of 1 if the average degree of obstacle is 3 or greater for both tax rates and tax administration, and 0 otherwise.

Control Variables in the Investment Efficiency Test

LogAssets = the log of total assets;

LogAge = the log of the age of the firm in years;

Tang = the ratio of property, plant, and equipment to total assets;

Slack = the ratio of cash to total assets; and

Audit = an indicator variable that takes on the value of 1 if annual financial statement is reviewed by an external auditor, and 0 otherwise.

Additional Variables Used in the Additional Tests

Collateral = the approximate value of collateral required as a percentage of the loan value;

Lev = financial leverage, measured as long-term liabilities divided by total assets;

BConIndex = the business constraint index, which is the average of indicator variables defined as equal to 1 if the firm perceives major or very severe obstacles in telecommunications, electricity, transportation, access to land, and skills and education of available workers, and 0 otherwise;

LConIndex = the legal constraint index, which is the average of indicator variables defined as equal to 1 if the firm perceives major or very severe obstacles in corruption, crime, and anticompetitive practices, and 0 otherwise;

(continued on next page)

APPENDIX A (continued)

PEConIndex = the political and economic index, which is the average of indicator variables defined as equal to 1 if the firm perceives major or very severe obstacles in customs regulations, labor regulations, business licensing, regulatory policy uncertainty, and the macroeconomic situation, and 0 otherwise;

AssetGrowth = asset growth, measured as the natural log of total assets in year $t-1$ divided by total assets in $t-2$;

StdRev = the standard deviation of sales in the past three years;

StdInvest = the standard deviation of investments in the past three years;

OperCycle = operating cycle calculated by the following formula: (average accounts receivables/sales) × 365 + (average inventory/cost of goods sold) × 365. We use the log of operating cycle for our tests;

Loss = an indicator variable referring to whether a firm reports a loss;

ROA = return on assets, measured as net income divided by lagged total assets;

Export = an indicator variable referring to whether the firm exports;

Foreign = an indicator variable referring to whether any foreign company or individual has a financial stake in the ownership of the firm;

FinCon = an indicator variable that takes on the value of 1 if the degree of obstacle for both access to financing and cost of financing is larger than 3, and 0 otherwise;

Control = an indicator variable that takes on the value of 1 if the largest shareholder owns more than 50 percent of the firm;

LocalMarket = the firm's reliance on local product markets, measured as the average of the percentage of domestic sales of outputs and the percentage of domestic purchases of inputs;

LegalOrigin = an indicator variable that equals 1 if the legal origin of the Company Law or Commercial Code of the sample country is English, and 0 otherwise (source: Djankov et al. 2007);

CreditorRights = an index aggregating rights of secured lenders that are defined in laws and regulations. The index ranges from 0 (weak creditor rights) to 4 (strong creditor rights) and is constructed as at January for every year from 1978 to 2003 (source: Djankov et al. 2007);

FinRegulation = an indicator variable that equals 1 if the sample firm belongs to a country that has substantial equivalent financial reporting regulation of private firms compared to public firms, and 0 otherwise; and

CashConstraint = An indicator variable that equals 1 if cash balance scaled by total assets is above the sample median, and 0 otherwise.

All data, unless otherwise noted, are from the World Bank's Enterprise Survey.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1289–1319

American Accounting Association
DOI: 10.2308/accr-10041

Monitoring by Auditors: The Case of Public Housing Authorities

Barbara Murray Grein
Drexel University

Stefanie L. Tate
University of Massachusetts Lowell

ABSTRACT: We take advantage of the unique reporting requirements of nonprofit public housing authorities (PHAs) to study the effect of audits on financial information both generally and when there are management incentives to misreport financial data. There is little prior research on the effect of audit adjustments in nonprofit settings and conflicting research on how auditors react to management's incentives to misreport. We identify potential financial statement areas at risk of manipulation based on incentives specific to public housing authorities. Using pre- and post-audit financial data for almost 3,600 PHAs across seven years, we find that auditors make economically and statistically significant adjustments to PHA financial statements. In addition, we find evidence that audits appear to reduce potential management bias, particularly to reduce risks of overstatement. Overall, audits appear to matter in this nonprofit, low-litigation risk setting where there is a large concentration of non-Big 4 auditors.

Keywords: *auditor monitoring; audit adjustments; earnings management; management incentives; public housing authorities; nonprofit organizations.*

Data Availability: *Data are available from public sources.*

I. INTRODUCTION

We take advantage of the unique reporting requirements of nonprofit public housing authorities (PHAs) to study the effect of audits on financial information both generally and when there are management incentives to misreport financial data. This research question is important for two reasons. First, while there is extensive research indicating that managers manipulate financial data to meet various incentives (e.g., Healy 1985; Hand 1989;

We thank two anonymous reviewers, Julia Higgs (discussant), Karen Kitching (discussant), Ranjani Krishnan (editor), Steven Kachelmeier (senior editor), workshop participants at the 2009 Government and Nonprofit Section Midyear Meeting, 2007 American Accounting Association Annual Meeting, 2007 Northeast Regional Meeting, and the University of Massachusetts for constructive criticisms and suggestions. We also thank Elizabeth Hanson and Nicholas Miele of the U.S. Department of Housing and Urban Development, as well as the managers and partners at the five public accounting firms we contacted for their insights. All remaining errors are of course our own.

Editor's note: Accepted by Ranjani Krishnan.

Submitted: August 2008
Accepted: January 2011
Published Online: April 2011

Burgstahler and Dichev 1997; Degeorge et al. 1999), and that auditors make significant adjustments as part of financial statement audits (e.g., Hylas and Ashton 1982; Kreutzfeldt and Wallace 1986; Kinney and McDaniel 1989; Icerman and Hillison 1991; Houghton and Fogarty 1991; Bell and Knechel 1994; Wright and Wright 1997), the evidence on how auditors react to management's incentives to misreport is mixed (e.g., Libby and Kinney 2000; Nelson et al. 2002; Ng 2007). Second, very few studies have focused on auditor adjustments in the nonprofit sector. Given that the nonprofit sector makes up more than 5 percent of the United States GDP,[1] and significant tax dollars are provided to public housing authorities specifically, we need a better understanding of how these organizations operate and are monitored.

We begin by analyzing the number and significance of audit adjustments by comparing pre- and post-audit data. We then identify potential financial statement areas at risk of manipulation based on incentives specific to public housing authorities. PHAs receive significant funding from the Department of Housing and Urban Development (HUD), which in turn monitors PHAs' efficiency, liquidity, and profitability. Better PHA operations lead to more funding opportunities and less intense oversight by HUD. Using pre- and post-audit financial data submitted by PHAs, we examine whether there is a differential effect on audit-related adjustments between PHAs with expected incentives to misreport financial data and PHAs with no such incentives. In addition, we consider which accounts are adjusted in response to PHA incentives using both univariate and multivariate analyses.

Using more than 19,000 observations across seven years for almost 3,600 nonprofit public housing authorities, we find that auditors make adjustments to PHA financial statements, a large number of which are material. These adjustments result in economically and statistically significant increasing as well as decreasing corrections across the financial statements. In addition, we find evidence that audits reduce potential management bias to meet HUD-established financial thresholds. While the evidence does not support audits reducing bias across all of the financial accounts or incentives tested, our results are consistent with audits focusing on overstatement risks.

Section II provides institutional background on public housing authorities. Section III provides the hypotheses development and Section IV discusses the data. Sections V, VI, and VII present our methodology and results on audit adjustments, reducing management bias, and account-specific analyses, respectively. Section VIII provides the results of sensitivity tests, and Section IX concludes.

II. INSTITUTIONAL BACKGROUND

Public housing authorities exist to assist the Department of Housing and Urban Development in meeting its mission "to increase home ownership, support community development and increase access to affordable housing free from discrimination," (http://www.hud.gov). PHAs, established by individual cities or counties as the need arises, own, maintain, and operate housing for low-income individuals and families. PHAs are ultimately responsible to HUD, but each city or county maintains a Board of Commissioners, whose responsibility is direct oversight of the PHA.

As recipients of significant HUD funding, PHAs are subject to HUD and auditor monitoring. All PHAs are required to file with HUD unaudited financial data within two months of fiscal year-end, and audited financial data within nine months of year-end. Financial data, both pre- and post-audit, are provided at the general ledger level, and is standardized across all PHAs. PHAs that receive more than $300,000 in federal funding are required to submit to an audit performed in accordance with the Office of Management and Budget's Circular No. A-133 (A-133 audit, Office of Management and Budget 1997).[2] The A-133 audit requires a specially trained auditor to provide an opinion on the financial statements, schedule of federal awards, and compliance with federal regulations.[3] In addition, the auditor must provide a report on the adequacy of internal controls and a summary of findings and questioned costs.[4]

---

[1] Statistics are for 2005 from "The Nonprofit Sector in Brief. Facts and Figures from the *Nonprofit Almanac 2007*," the Urban Institute, http://www.urban.org/UploadedPDF/311373_nonprofit_sector.pdf (accessed June 2009); 2005 is the most recent year for which data are available.

HUD evaluates each PHA annually on four factors: financial (30 points), physical (30 points), management (30 points), and resident (10 points), summarized in a 100-point Public Housing Assessment System (PHAS) score.[5] The PHAS score is used to classify a PHA as a high performer, standard performer, or troubled performer. A PHA is classified as a high performer if its PHAS score is 90 or above and it receives at least 60 percent on all four factors. A standard performer has a score between 60 and 90 and receives at least 60 percent on the financial, physical, and management factors. All other PHAs are considered troubled. High performers are eligible for capital fund bonuses and biannual, rather than annual, physical inspections. Troubled PHAs are more closely supervised by the HUD field office, and continual troubled ratings can result in being referred to the HUD Enforcement Center and eventually suspended or disbarred (http://www.hud.gov/utilities/intercept.cfm?/offices/reac/products/PDFs/PHAS_NASS.ppt).

We focus on the financial portion of the PHAS score, the Financial Assessment Subsystem (FASS) score, which is based on the PHA's performance on the ratios shown in Exhibit 1.

Appendix A provides details on the calculations of each of the ratios. We discuss the use of and point allocations for all ratios with the exception of occupancy loss in Section VI. Occupancy loss measures the percent of apartments not leased during the year. While an important measure of PHA performance, it is not based on financial information subject to management bias or audit. As such we exclude this ratio from our subsequent discussions.

III. HYPOTHESES DEVELOPMENT

Jensen and Meckling (1976) suggest that monitoring by an independent third-party auditor is a means of reducing agency costs.[6] In support of this critical role, Watts and Zimmerman (1983) find that audits have been around in some form since the early development of corporations. We examine the effect of audits on financial information generally as well as whether audits reduce management bias related to achieving financial targets.

Audit Adjustments

The primary means auditors have to reduce agency costs is by reducing financial statement misstatements. Professional standards require auditors to assess the risks of material misstatements due to error and fraud, and plan and perform their audit accordingly to detect them (AICPA 1972,

---

[2] The limit was raised to $500,000 in June 2003.

[3] Auditors performing A-133 audits are required to receive training specific to A-133 and/or federal government Yellow Book audits.

[4] Some PHAs not subject to an A-133 audit may be required by their associated local government (e.g., city or county) to have a Yellow Book audit. These audits are similar to an A-133 audit and include an assessment of internal controls and compliance with federal regulations in addition to the financial statement audit.

[5] The financial factor score is based on the PHA's financial performance in six areas detailed later in this study. The physical component is based on results of an inspection of the PHA property for more than 500 conditions. The management component is scored based on results of a PHA certification of its management performance over six items including vacant unit turnaround time, security, and economic self-sufficiency. Finally, the resident factor score is based on an annual survey of resident satisfaction (http://www.hud.gov/offices/reac/atglance.cfm, accessed June 2010).

[6] Jensen and Meckling (1976) define agency costs as monitoring expenditures by the principal, bonding agents by the principal, and loss in principal's welfare due to the agents not acting in the principal's best interest.

Exhibit 1

Financial Ratio Point Values

	Financial Score Component	Maximum Points
CR	Current ratio	9.0
ME	Number of months expendable fund balance	9.0
TRO	Tenant receivables outstanding	4.5
OL	Occupancy loss	4.5
UTIL	Expense management/utility consumption	1.5
NIFB	Net income as a percent of expendable fund balance	1.5
	Total	30.0

2002, 2006a, 2006b).[7] Prior research has consistently shown that auditors make material adjustments to financial statements (e.g., Hylas and Ashton 1982; Kreutzfeldt and Wallace 1986; Kinney and McDaniel 1989; Icerman and Hillison 1991; Houghton and Fogarty 1991; Bell and Knechel 1994; Wright and Wright 1997). In a meta-type analysis examining 13 studies including those listed above, Kinney and Martin (1994, 154) find that "60 to 90 percent of the audits have non-zero adjustments and about 20 to 60 percent have 'material' aggregate adjustments." We propose the following hypothesis, consistent with professional standards and prior research:

H1: Audits result in significant adjustments to PHA financial statements, *ceteris paribus.*

There are several reasons why we cannot be certain prior results will generalize to our setting. First, several studies find audit-related adjustments vary in number and magnitude across industries (e.g., Hylas and Ashton 1982; Kreutzfeldt and Wallace 1986; Houghton and Fogarty 1991), and only a minority of audit-adjustment studies include nonprofit organizations as part of their sample. Second, nonprofit entities in general and PHAs in particular represent low-litigation risk for auditors and the threat of litigation and associated economic loss is generally considered one motive for an audit firm to perform a high-quality audit.[8] Finally, prior audit adjustment research focuses almost exclusively on Big N auditors (see a review by Kinney and Martin [1994]), while PHAs are largely served by smaller audit firms frequently classified as lower quality. However, small audit firms are subject to the same AICPA rules as their larger counterparts. In addition, PHA audits in accordance with A-133 require specially trained auditors and are subject to quality control desk reviews by HUD. Major inadequacies or repetitive substandard performance by an auditor will be referred to the state licensing board, (A-133, Subpart D, Section 400, (a) (3) and (a) (5)). Further, DeAngelo (1981), among others, suggests that auditors perform audits at one quality level, implying that a high-quality auditor will maintain that same quality level in their PHA audits. Overall, it is an empirical question whether nonprofit PHA audits result in significant adjustments to the financial statements.

Reducing Management Bias

Extensive prior research indicates that management responds to incentives to "adjust" financial statements to maximize personal and firm wealth (see Healy and Wahlen [1999] for a summary). While nonprofit entities do not have incentives related to capital market pressures, they have incentives related to contracts and regulatory monitoring. A growing body of research finds that nonprofit managers respond to financial incentives (e.g., Duggan 2000) and, in particular, respond to incentives to manipulate financial information (e.g., Yetman 2001; Leone and Van Horn 2005; Krishnan et al. 2006; Jones and Roberts 2006; Keating et al. 2008).

---

[7] Audits of PHA financial statements are performed under the same audit standards as for-profit corporations.

[8] Consistent with DeAngelo (1981), we define audit quality as the probability that an audit firm will detect and report material misstatements in the financial statements.

PHAs have incentives to attain or maintain a high PHAS score in order to maintain access to capital from HUD and to avoid costly additional regulatory monitoring. As part of the PHAS score, PHAs have an incentive to attain or maintain a high FASS score.[9] As a result, PHAs' *financial* incentives focus on meeting or beating pre-established thresholds for the six ratios that go into the FASS score. In terms of PHAs' incentives to manage financial information, it is important to note that HUD assesses the FASS score using *pre-audit* financial information. HUD penalizes a PHA by reducing their score if there are large differences between pre- and post-audit financial data. Discussions with external auditors and representatives at HUD confirm the existence of PHA financial incentives.[10] Both groups deemed the incentive to avoid being a troubled performer and the resulting additional oversight by HUD as high. They also perceived a strong incentive to reduce HUD oversight and be eligible for additional funding by qualifying as a high performer. These incentives represent regulatory motivations and are similar to monitoring via debt covenants, both of which prior research has found are associated with accounting discretion (e.g., for regulatory incentives see Scholes et al. 1990; Jones 1991; Petroni 1992; Beatty et al. 1995; Collins et al. 1995; for debt covenants see Watts and Zimmerman 1986; DeFond and Jiambalvo 1994; Sweeney 1994.

Potential financial statement misreporting naturally brings up the question, "Where are the auditors?" Professional standards require auditors to assess the risks of material misstatements and plan and perform their audits to detect such misstatements. There is an extensive body of research on the role of the auditor in mitigating misstatements. The research most directly related to this study examines auditors' (1) ability to identify management incentives and (2) willingness to require adjusting entries when identified misstatements are consistent with management's incentives to manipulate financial reporting.[11] Experimental research has found auditors successfully identify situations in which management has incentives to manage financial information (e.g., Hirst 1994; Libby and Kinney 2000; Ng 2007). Whether auditors require adjusting entries in the face of management incentives is less clear. Libby and Kinney (2000) find auditors are *less* likely to require an adjustment when the adjustment would result in management *not* meeting analyst forecasts. Ng (2007) finds that auditors' likelihood of proposing a quantitatively immaterial but qualitatively material adjustment affecting a client's ability to meet earnings thresholds ranges from 8 percent related to an analyst forecast threshold to 70 percent for a prior year earnings threshold.[12] Using survey data, Nelson et al. (2002) find that auditors are more likely to adjust material earnings management attempts.

---

[9] As discussed above, the FASS score represents up to 30 points of the 100-point PHAS score. Further, PHAs must earn at least 60 percent, or 18 points, on the financial portion of the PHAS score in order to avoid being classified as troubled.

[10] We contacted partners and/or managers at five of the top 20 audit firms within our sample (ranked by percent of regional assets audited) to discuss PHA audit risks, accounting sophistication and complexity, and manager incentives to meet HUD ratios. The firms contacted audited a total of 402 observations for 103 different clients within our database. These firms audit a wide range of PHAs with total audited assets (revenue) ranging from $149,147 to $5.23 billion ($290,622 to $345 million), and average audited assets (revenue) of $105 million ($21.4 million).

[11] A related stream of research investigates the association between auditors and earnings management, (e.g., Becker et al. 1998; Kim et al. 2003; Balsam et al. 2003; Chung et al. 2005). This stream of research generally finds high-quality auditors are associated with clients with lower post-audit earnings management.

[12] The 8 percent finding was related to a *subjective* audit difference affecting the client's ability to meet an analyst forecast threshold and the 70 percent finding was for an *objective* audit difference affecting the client's ability to meet a prior-year earnings threshold when the auditor received explicit materiality guidance.

Based on our discussions with partners and managers on PHA audit engagements, prior research, and auditing standards, we expect auditors to be aware of PHA incentives and to plan and perform audits to identify and correct material misstatements that could result from managerial manipulation of the financial data. We test the following hypothesis:

> **H2:** Audits reduce potential management bias when PHAs have greater incentives to meet financial targets, *ceteris paribus.*

IV. DATA AND DESCRIPTIVE STATISTICS

Data

The data used in this study were obtained from HUD for all nonprofit public housing authorities.[13] Although PHAs submitted relevant data to HUD starting in 1999, the PHAS scoring system was not enforced until fiscal years ending on or after September 2001. Accordingly, our analyses are performed using data for fiscal years ending between September 2001 and December 2007. We eliminated 161 single-year observations and 120 non-consecutive observations because we require lagged values for various analyses. Our final sample includes 19,569 observations for 3,567 PHAs.

The dataset includes pre- and post-audit financial data at the general ledger level and A-133/Yellow Book audit results (e.g., findings, questioned costs, and opinions on internal controls and compliance with federal regulations). In addition, the dataset includes total PHAS and FASS scores for some, but not all, observations. Where the FASS scores were missing, we estimate FASS scores based on the individual values of each of the ratios. Our analyses and results include both the actual and estimated FASS scores.

Descriptive Statistics

In Table 1, we provide descriptive information for the PHAs in our sample. Given that this is a relatively understudied industry, we provide analyses of variables we believe help to better understand the industry—both the PHAs themselves in Panel A and the PHA auditors in Panel B. The average PHA has almost $19 million in assets, with more than $5 million in current assets and more than $11 million in net property, plant, and equipment. The median PHA is much smaller, with total assets just under $3.5 million and net property, plant, and equipment just over $2.6 million. Mean (median) total liabilities are about $6 million ($165,000), and mean (median) current liabilities are about $1.8 million ($112,000). Mean net income is approximately $313,000, while the median is only $20,000, consistent with the PHAs' nonprofit status.

Table 1, Panel B provides descriptive data on PHA auditors. There are 770 different firms auditing PHAs in our sample, with four Big 5 auditors, and 28 of the Top 50 auditors.[14] Sample firms audit a mean (median) of 25 (5) client-years with mean (median) assets of $41.6 million ($5.0 million). Big 5 and Top 50 auditors audit 0.6 percent and 3.4 percent of our sample, respectively, and while they audit about the same number of client-years on average, Big 5 clients are much larger. The average PHA audited by the Big 5 has $872 million in total assets while the average Top 50-audited PHA has $210 million in total assets.

---

[13] While PHAs can be for-profit or nonprofit, only nonprofit financial data are publicly available from HUD.

[14] Although Arthur Andersen was still in existence at the beginning of our sample period and they did perform some PHA audits in the past, they did not perform any PHA audits during our sample period. We use the annual rankings published in *Accounting Today* for each of our sample years to determine the Top 50 firms.

TABLE 1

Descriptive Statistics for 3,567 Public Housing Authorities for the Period September 2001 through December 2007

Panel A: Public Housing Authority Post-Audit Descriptive Data

Variable	Mean	Median	Quartile 1	Quartile 3	Standard Deviation
Assets	18,900,000	3,458,861	1,305,485	8,505,840	176,000,000
Current Assets	5,216,077	609,413	228,464	1,760,747	58,500,000
Net Property, Plant, and Equipment	11,200,000	2,651,548	974,607	6,461,676	103,000,000
Liabilities	5,996,448	165,195	44,027	939,515	87,500,000
Current Liabilities	1,761,780	112,369	36,548	417,879	25,900,000
Non-Current Liabilities	4,234,668	15,135	0	335,512	70,200,000
Revenue	7,917,978	1,211,460	424,502	3,918,896	60,500,000
Net Income	313,536	20,471	−54,593	165,993	9,592,284

Panel B: Auditors of Public Housing Authorities Descriptive Data

	All Auditors	Big 5	Top 50
Percentage of observations audited	100%	0.6%	3.4%
Mean number of audits per firm	25	29	24
Median number of audits per firm	5	23	15
Mean assets audited per firm	41,600,000	872,000,000	210,000,000
Median assets audited per firm	5,008,311	892,000,000	68,600,000
Total number of different firms	770	4[a]	28

---

[a] Although Arthur Andersen was still in existence at the beginning of our sample period and they did perform some PHA audits in the past, they did not perform any PHA audits during our sample period.

V. AUDIT ADJUSTMENTS

Research Design

If monitoring by auditors reduces misstatements in financial information, then to the extent that there are significant misstatements in the pre-audit financial information, we expect the audit to result in adjustments and the post-audit financial information to differ significantly from the pre-audit information. To test H1, we examine the distribution of audit adjustments (defined as the audited value less the unaudited value) for major financial statement categories used in calculating the FASS ratios, as well as the ratios themselves. Statistically significant differences between the pre- and post-audit balances would suggest that, on average, auditors make adjustments to the financial statements. Next, we examine the direction of audit adjustments by evaluating the number and percentage of audits that result in increases, decreases, or no adjustment for each financial statement category and ratio. Finally, we examine whether the changes from pre- to post-audit are economically significant by evaluating the number and percentage of observations that have material changes to net income, total assets, or both. We consider two potential materiality thresholds based on pre-audit amounts: one-half percent of total assets and one-half percent of total revenue.[15] The actual materiality thresholds used in the individual audit engagements are unavailable.

Results

Table 2, Panel A provides the means, medians, and 1st and 99th percentiles of audit-related adjustments to the overall ratios and the accounts that make up the ratios. We analyze unscaled adjustments to both and adjustments scaled by total assets for the specific accounts.[16] While the means are relatively small in magnitude, the ranges of adjustments are large. For example, adjustments to current liabilities are an average increase of $7,911 and range from a decrease of almost $1 million to a $1.3 million increase, or a decrease of almost 7 percent of total assets to an increase of more than 7 percent of total assets. The mean change to total expenses is $17,234, ranging from a decrease of $0.6 million to an increase of $0.9 million, or a decrease of almost 5 percent of total assets to an increase of more than 8 percent of total assets. All scaled and unscaled adjustments are statistically significantly different from zero ($p < 0.05$, untabulated) with the exception of adjustments in tenant receivables outstanding (*TRO*) and the related accounts tenant receivables and tenant revenue.

Table 2, Panel B presents a summary of the number and percent of audits that result in increases, decreases, or no adjustments to financial statement accounts and ratios. The percentage of observations with no adjustment ranges from a low of 31 percent for net income expendable fund balance to a high of 85 percent for tenant receivables.[17] The current ratio, months expendable fund balance, other accounts receivable, and tenant revenue had a greater incidence of ratio/account-decreasing adjustments than ratio/account-increasing adjustments. All other accounts have a greater incidence of ratio/account-increasing than -decreasing adjustments.

Table 2, Panel C presents the number and percent of all observations that have material adjustments to net income, total assets, and both net income and total assets. Depending on the materiality threshold used, between 28 and 39 percent of all observations have material adjustments to net income or total assets, which is within the range found in prior research.[18] Twenty to 29 percent of observations have material adjustments to both net income and assets.

The results in Table 2 support H1 suggesting that audits result in material adjustments to the financial statements in a significant percentage of engagements in this low-litigation risk setting with a large concentration of non-Big 4 auditors. We acknowledge that our tests may overstate the effect of the audit on the financial statements, given the potential for a long time lag and subsequent events between pre- and post-audit data. However, discussions with PHA auditors suggest that most adjustments are not the result of additional information becoming available between the two filings; instead, adjustments are made as a result of improper application of GAAP and improper or inadequate closing procedures. Alternatively, examination of audit adjustments could understate the effect of audits. Kinney and Martin (1994) suggest three reasons that apply here. First, client firms anticipating an audit are likely to be more careful in maintaining their accounting records and less

---

[15] These materiality thresholds are consistent with those used in prior research on audit adjustments (e.g., Kinney and Martin 1994; Wright and Wright 1997). Further, Messier et al. (2006) recommend using one-half percent of total revenues or total expenses for nonprofit entities.

[16] To mitigate the effects of extreme observations, we winsorize the data at the 1st and 99th percentiles.

[17] Kinney (1979), investigating audit-related adjustments, and Kreutzfeldt and Wallace (1986), investigating waived and booked audit adjustments, document financial statement accounts with a similar range of no adjustments.

[18] Wright and Ashton (1989) find 73.7 percent of sample firms had material proposed and/or booked adjustments and Bell et al. (1992) as quoted by Kinney and Martin (1994) find 22.7 percent of sample firms had material adjustments.

TABLE 2
Analysis of Audit-Related Adjustments

Panel A: Adjustments from Pre- to Post-Audit

	Adjustment[a]		Adjustment as a Percent of Assets[a]	
Variable[b]	**Mean (Median)**	**1% (99%)**	**Mean (Median)**	**1% (99%)**
Current Ratio (*CR*)	−0.20	−14	NA	NA
	(0)	(10)		
Months Exp. Fund Bal. (*ME*)	−0.15	−10	NA	NA
	(0)	(8)		
Tenant Rec. Outstand. (*TRO*)	−0.01	−11	NA	NA
	(0)	(8)		
Utility Consumption (*UTIL*)	0.33	−20	NA	NA
	(0)	(29)		
Net Inc. Exp. Fund Bal. (*NIFB*)	0.02	−1.35	NA	NA
	(0)	(1.83)		
Tenant Accts. Rec. (*TENAR*)	−15	−17,578	0.00%	−0.17%
	(0)	(16,200)	(0.00%)	(0.13%)
Other Current Assets (*OCASS*)	−4,900	−1,565,780	−0.04%	−9.85%
	(0)	(1,511,014)	(0.00%)	(8.51%)
Current Liabilities (*CLIAB*)	7,911	−1,307,728	0.10%	−7.97%
	(0)	(1,549,023)	(0.00%)	(9.18%)
Tenant Revenue (*TENREV*)	267	−146,671	−0.01%	−1.45%
	(0)	(157,556)	(0.00%)	(1.41%)
Other Revenue (*OREV*)	12,286	−1,022,886	0.14%	−7.98%
	(0)	(1,355,848)	(0.00%)	(11.35%)
Expenses (*EXP*)	17,234	−692,013	0.17%	−6.16%
	(0)	(1,176,730)	(0.00%)	(9.81%)

Panel B: Number and Percent of Observations of No Adjustment, Account Increase, and Account Decrease from Pre- to Post-Audit

	No Adjustment		Account Increase		Account Decrease	
Variable[b]	**Number of Obs.**	**Percent of Obs.[c]**	**Number of Obs.**	**Percent of Obs.[c]**	**Number of Obs.**	**Percent of Obs.[c]**
Current Ratio (*CR*)	7,082	36.19%	5,568	28.45%	6,805	34.77%
Months Exp. Fund Bal. (*ME*)	6,135	31.35%	6,205	31.71%	7,225	36.92%
Tenant Rec. Outstanc. (*TRO*)	13,416	68.56%	3,029	15.48%	2,171	11.09%
Utility Consumption (*UTIL*)	7,668	39.18%	6,573	33.59%	5,306	27.11%
Net Inc. Exp. Fund Bal. (*NIFB*)	6,017	30.75%	7,213	36.86%	6,294	32.16%
Tenant Accts. Rec. (*TENAR*)	16,626	84.96%	1,696	8.67%	1,247	6.37%
Other Current Assets (*OCASS*)	10,224	52.25%	4,514	23.07%	4,831	24.69%
Current Liabilities (*CLIAB*)	8,636	44.13%	6,108	31.21%	4,825	24.66%
Tenant Revenue (*TENREV*)	14,914	76.22%	2,026	10.35%	2,629	13.43%
Other Revenue (*OREV*)	9,122	46.61%	6,081	31.07%	4,366	22.31%
Expenses (*EXP*)	7,949	40.62%	6,709	34.28%	4,911	25.10%

(continued on next page)

TABLE 2 (continued)

Panel C: Number and Percent of Material Auditor-Corrected Misstatements

	$\lvert\Delta$Net Income$\rvert$		$\lvert\Delta$Asset$\rvert$		$\lvert\Delta$Net Income$\rvert$ and $\lvert\Delta$Asset$\rvert$	
Greater Than	**Number of Obs.**	**Percent of Obs.**[c]	**Number of Obs.**	**Percent of Obs.**[c]	**Number of Obs.**	**Percent of Obs.**[c]
½% × Pre-Audit Assets	5,487	28.04%	5,995	30.64%	3,959	20.23%
½% × Pre-Audit Revenue	7,721	39.46%	7,550	38.58%	5,732	29.29%

[a] Adjustment is calculated for each observation as audited value − unaudited value. Median, 1st, and 99th percentile values are from the raw data. Reported means are based on data that were winsorized at the 1st and 99th percentiles.
[b] Variables defined in Appendix B.
[c] Percent of observations calculated as the number of observations divided by 19,569, the total number of observations from September 2001 through December 2007.

likely to bias amounts involving managerial discretion or judgment. In our setting, this anticipation is reinforced by the fact that both pre- and post-audit data are examined by HUD, and PHAs can incur penalties for large differences between the two sets of numbers. Second, our approach does not include misstatements prevented because the client consulted with the auditor prior to recording the transaction. Finally, our approach does not include misstatements that were avoided due to client controls in place to reduce audit effort and fees. It is unclear whether and in what direction, on average, these factors would bias our results.

VI. REDUCING MANAGEMENT BIAS

H2 predicts that audits reduce potential management bias when the client has expected incentives to meet financial targets. We predict audit adjustments are in the opposite direction of a PHA's incentives. We expect PHAs close to classification or point thresholds to have the greatest incentives to report opportunistically (Healy 1985; Burgstahler and Dichev 1997; Degeorge et al. 1999). With this in mind, we identify two categories of incentives—those associated with avoiding troubled performer classification and those associated with maximizing points on individual ratio scores.[19] We then examine whether there is a difference in audit-related adjustments on FASS ratios between PHAs with expected incentives to manipulate financial data and PHAs without such incentives. Next, we consider which accounts are adjusted in response to PHA incentives to misreport.

Research Design—Troubled Performer Incentives

We identify and examine three sets of PHAs as having had greater incentives to bias financial data related to the troubled performer classification: (1) PHAs with pre-audit PHAS scores at or just above 60 points (*P60-Up* = 1, 0 otherwise), (2) PHAs with pre-audit FASS scores at or just above 18 points (*F-Up* = 1, 0 otherwise), and (3) PHAs with pre-audit FASS scores in the lowest two percentiles under 18 points (*F-Down* = 1, 0 otherwise).[20] "Just above" is defined as the two percentiles of PHAs that fall above 60 (18) after ranking all PHAs by their PHAS (FASS) score. PHAs in the first two categories, *P60-Up* = 1 and *F-Up* = 1, are expected to have had greater incentives to increase their financial score to avoid the troubled performer classification. We maintain the assumption that PHAs with ratio values just missing the thresholds had an incentive to misreport but insufficient financial flexibility and, therefore, likely did not engage in misreporting. In sensitivity analyses we relax this assumption. PHAs in the last category, *F-Down* = 1, are expected to have had greater incentives to decrease their score in the current year to create financial slack (cookie jar reserves) for use in the future.

---

[19] We also consider PHA incentives to be classified as a high performer by having a PHAS score of 90 points or better. The results of all analyses related to this incentive are not statistically significant. Similar to the findings in Duggan (2000), this may be associated with the potential of the local government to cut a PHA's funding if they receive additional funds from HUD, mitigating one of the benefits of being classified as a high performer.

[20] Avoiding the troubled classification is based on achieving a PHAS score of at least 60 points and achieving minimum thresholds (60 percent or 18 points) on the financial, physical, and management portions. Consistent with HUD scoring, all incentives are coded using pre-audit data. Percentiles are determined by first ranking all PHAs by FASS score. We use the 1st and 2nd percentiles with FASS scores below 18 for category (3).

For each incentive category, we compare audit-related changes in the FASS score component ratios between PHAs with incentives and those without incentives. If audits reduce misreporting, we expect adjustments to be greater (in the predicted direction) for PHAs with incentives to bias financial data than for PHAs without incentives. The scoring for the current ratio, months expendable fund balance, and net income expendable fund balance is such that, in general, higher values of the ratios result in more points awarded. As such, we expect audit adjustments to these ratios for *P60-Up-* and *F-Up*-incentive PHAs to be more negative/less positive than for no-incentive PHAs. The reverse is true for *F-Down*-incentive PHAs. The scoring for tenant receivables outstanding and utility consumption is such that lower values result in more points awarded. Therefore, we expect more positive/less negative adjustments for *P60-Up-* and *F-Up*-incentive PHAs as compared to no-incentive PHAs, and the reverse for *F-Down*-incentive PHAs. Details of the ratio scoring are provided under "Research Design—Individual Ratio Incentives," below.

Results—Troubled Performer Incentives

Table 3 presents the results of our tests. The results in Panel A for PHAs identified as having greater incentives to increase their PHAS score (*P60-Up* = 1) indicate that, on average, audits result in adjustments that reduce the current ratio and months expendable fund balance and make adjustments to increase utility consumption, consistent with our expectations. Contrary to expectations, the mean adjustments to net income expendable fund balance are more positive for *P60-Up*-incentive PHAs than no-incentive PHAs. When PHAs have an incentive to increase their FASS score to avoid a troubled designation (*F-Up* = 1), auditors only make statistically significant adjustments to months expendable fund balance, as indicated in Panel B.

Table 3, Panel C presents the results comparing PHAs with an incentive to create financial slack as it relates to their FASS score (*F-Down* = 1) to those with no incentive. The significantly positive difference between *F-Down*-incentive and no-incentive PHAs for the current ratio, months expendable fund balance, and net income expendable fund balance are consistent with auditors attempting to constrain management bias. However, the positive significant difference for utility consumption is inconsistent with expectations. While not all of the findings are consistent with expectations, the overall results in Table 3 are generally consistent with audits reducing bias, supporting H2. Further, the results are consistent with auditors concentrating on the ratios that can have the most impact; the current ratio and months expendable fund balance are each worth 9 points, at least double the points of the other ratios.

Research Design—Individual Ratio Incentives

Next, we identify PHAs with greater incentives to bias financial reporting based on point allocations for the individual FASS score ratios. This approach relaxes the previous case and assumes that, all else equal, PHAs like to report higher ratio, FASS, and PHAS scores. Based on HUD's scoring for each ratio, we categorize PHAs into those with an apparent incentive to increase

TABLE 3

Comparison of Mean Changes in Ratios between PHAs with Incentives to Increase or Decrease PHAS or FASS Scores and All Other PHAs

Panel A: Adjustments When Management Has Incentives to *Increase* PHAS Score[a]

					Mean Adjustment[b]			
Ratio[c]	**Pts.[d]**	**Exp.[d]**	**Incentive n**	**No Incentive n**	**Incentive PHAs**	**No Incentive PHAs**	**Diff.[d]**	**Sig.**
Current Ratio	9	–	330	13,114	–0.28	–0.10	–0.18	**
Months Exp Fund Bal.	9	–	330	13,114	–0.45	–0.09	–0.36	***
Tenant Rec. Outstand.	4.5	+	330	13,114	–0.18	–0.01	–0.17	
Utility Consumption	4.5	+	330	13,114	0.93	0.30	0.62	**
Net Inc. Exp. Fund Bal.	1.5	–	330	13,114	0.13	0.01	0.12	###

Panel B: Adjustments When Management Has Incentives to *Increase* FASS Score[e]

					Mean Adjustment[b]			
Ratio[c]	**Pts.[d]**	**Exp.[d]**	**Incentive n**	**No Incentive n**	**Incentive PHAs**	**No Incentive PHAs**	**Diff.[d]**	**Sig.**
Current Ratio	9	–	473	18,718	–0.26	–0.21	–0.05	
Months Exp. Fund Bal.	9	–	473	18,718	0.02	–0.17	0.20	**
Tenant Rec. Outstand.	4.5	+	473	18,718	0.07	–0.02	0.09	
Utility Consumption	4.5	+	473	18,718	0.27	0.32	–0.05	
Net Inc. Exp. Fund Bal.	1.5	–	473	18,718	0.03	0.01	0.02	

Panel C: Adjustments When Management Has Incentives to *Decrease* FASS Score[f]

					Mean Adjustment[b]			
Ratio[c]	**Pts.[d]**	**Exp.[d]**	**Incentive n**	**No Incentive n**	**Incentive PHAs**	**No Incentive PHAs**	**Diff.[d]**	**Sig.**
Current Ratio	9	+	375	18,718	0.38	–0.21	0.59	***
Months Exp. Fund Bal.	9	+	375	18,718	1.05	–0.17	1.22	***
Tenant Rec. Outstand.	4.5	–	375	18,718	–0.05	–0.02	–0.03	
Utility Consumption	4.5	–	375	18,718	1.06	0.32	0.74	###
Net Inc. Exp. Fund Bal.	1.5	+	375	18,718	0.22	0.01	0.21	***

***, ** Change for PHAs with incentives is significantly different from PHAs with no incentives at p-value < 0.01 and 0.05, respectively. Results reported are one-tailed.

Change for PHAs with incentives is significantly different but in the unexpected direction from PHAs with no incentives at p-value < 0.01. Results reported are two-tailed.

[a] Panel A compares PHAs with a strong incentive to increase their PHAS score to PHAs with no incentives related to their PHAS scores. PHAs with a PHAS score of less than 60 are considered troubled. Therefore, PHAs just earning 60 points and the two percentiles above 60 are considered to be the most likely to have manipulated their financial data to obtain that score.

[b] t-tests are calculated using the mean change in the ratio.

[c] Ratio formulas are included in Appendix A.

[d] The points represent the maximum possible score for the ratio. The expected sign of the difference is based on our belief that audits reduce manipulation of data through adjusting entries. The differences in the mean changes between incentive PHAs and no-incentive PHAs are calculated from original, unrounded data and are compared to zero.

(continued on next page)

TABLE 3 (continued)

[e] Panel B compares PHAs with a strong incentive to increase their FASS score to PHAs with no incentives related to their FASS scores. PHAs with a FASS score of less than 18 are considered troubled. Therefore, PHAs just earning 18 points and the two percentiles above 18 are considered to be the most likely to have manipulated their financial data to obtain that score.

[f] Panel C compares PHAs with a strong incentive to decrease their FASS score to PHAs with no incentives related to their FASS scores. PHAs with a FASS score of less than 18 are considered troubled. Therefore, PHAs with the lowest FASS scores below 18, measured as in the 1st and 2nd percentiles of FASS scores, may have manipulated the financial data to create slack in the current year that may help them increase their FASS score in future years when accruals reverse.

the ratio, decrease the ratio, or no apparent incentive either way.[21] The following describes the scoring as well as our identification of corresponding incentives to bias financial reporting.

Current Ratio (CR) and Months Expendable Fund Balance (*ME*)

The current ratio and months expendable fund balance measure the PHA's ability to pay current liabilities and monthly expenses, respectively. Because the ratios have similar point allocations, we discuss them jointly. For fiscal years ending between September 2001 and June 2003, all PHAs with a ratio of 1.0 or greater earned 9 points; all other PHAs earned no points. In this period, PHAs with a current ratio (months expendable fund balance ratio) value equal to 1.0 and "just above" 1.0 are expected to have an incentive to manipulate their ratio up to maximize points and are coded as $CR\text{-}Up = 1$ ($ME\text{-}Up = 1$).[22]

For fiscal years ending after June 2003, point allocations are determined using the curve shown in Figure 1. PHAs receive no current ratio or months expendable fund balance points if the ratio value is less than 1.0. Points are awarded for ratios equal to 1.0 and points gradually increase to 9 for ratios between 1.0 and T_1. All PHAs with ratios between T_1 and T_2 earn 9 points, but points gradually decrease from 9 to 7.5 for values above T_2 but below T_3. All PHAs above T_3 earn 7.5 points. The threshold values vary by PHA size, were provided to all PHAs, and were published in Federal Registers. For this period, PHAs with a ratio value between 1.0 and T_1 and PHAs in the first two percentiles above T_1 are identified as having an incentive to manipulate their ratio up to increase their points ($CR\text{-}Up = 1$, $ME\text{-}Up = 1$). PHAs with ratio values between T_2 and T_3 have an incentive to decrease their ratio to increase points. To better focus the analyses, we exclude the latter PHAs from our primary analyses.[23]

PHAs with no ability to meet the 1.0 threshold (i.e., those with very low ratio values) could attempt to decrease their ratio in the current year to improve their chances of meeting the 1.0 threshold in the next year when accruals reverse. Accordingly, we identify PHAs in the lowest two percentiles (ensuring they were below a value of 1.0) as having had incentives to decrease their current (months expendable fund balance) ratio, coding them as $CR\text{-}Down = 1$ ($ME\text{-}Down = 1$). PHAs not identified as having had a greater incentive to increase or decrease their ratios are considered to have no incentives with respect to these ratios and are coded $CR\text{-}Up = 0$, $CR\text{-}Down = 0$, $ME\text{-}Up = 0$, and $ME\text{-}Down = 0$, respectively.

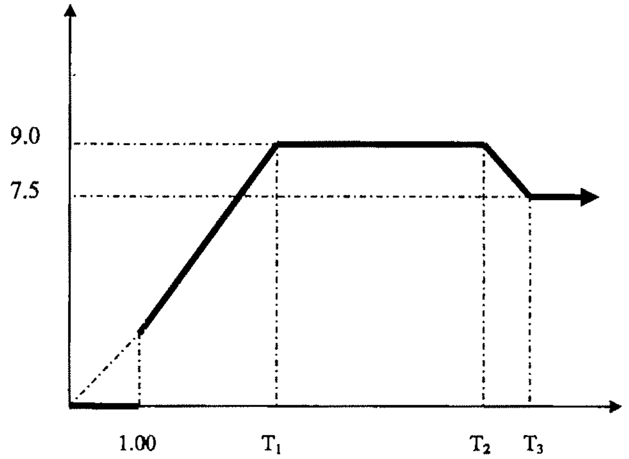
---

[21] We acknowledge some PHAs may not have an incentive to manipulate their ratio scores if it will not affect their performance category (high, standard, or troubled). Therefore, this approach could overstate the number of incentive-PHAs, resulting in weaker tests of our hypothesis.

[22] "Just above" is determined by ranking all PHAs by their *CR* (*ME*) value. PHAs in the two percentiles that fall above the threshold are "just above." Percentiles around other ratio thresholds were determined similarly.

[23] Supplemental tests of these PHAs indicate generally insignificant differences between audit adjustments of no-incentive PHAs and PHAs with an incentive to decrease their *CR* and *ME* ratios in order to increase their FASS score.

FIGURE 1
Point Allocation Graph for Current Ratio and Months Expendable Fund Balance for Fiscal Years Ended after June 2003



Tenant Receivables Outstanding (TRO)

Tenant receivables outstanding (*TRO*) is a measure of how efficiently the PHA collects tenant receivables; the point allocation for it is depicted in Figure 2. PHAs earn the maximum 4.5 points if they are below T_1, gradually decreasing points between T_1 and T_2, and earn no points if they are above T_2. Similar to the current ratio and months expendable fund balance, threshold values vary by PHA size category.

We categorize PHAs with *TRO* values between T_1 and T_2 and in the two percentiles below T_1 as having greater incentives to decrease their ratio to maximize their points (*TRO-Down* = 1). We identify PHAs with the lowest *TRO* values (in the first two percentiles) as having an incentive to increase their ratio with little or no risk of losing *TRO* points (*TRO-Up* = 1).[24] Increasing the ratio in the current year through accruals could help to decrease the ratio in future years when those accruals reverse, potentially helping the PHA to earn more points in the future. All other PHAs are considered to have no incentives with respect to the *TRO* ratio (*TRO-Up* = 0 and *TRO-Down* = 0).

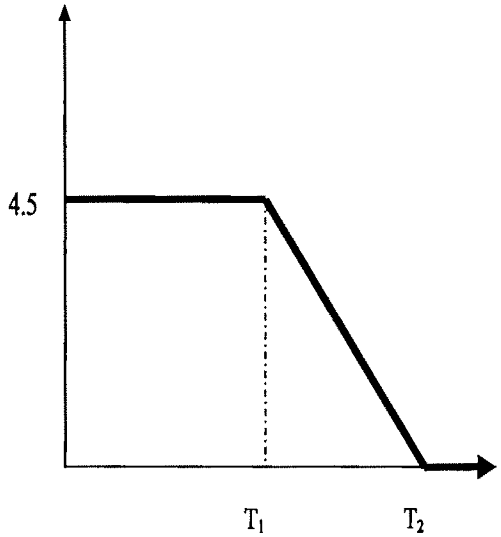
Expense Management/Utility Consumption (UTIL)

Expense management/utility consumption (*UTIL*) is a measure of how efficiently the PHA manages the costs per apartment. PHAs receive 1.5 points if *UTIL* is below T_1 for their region and size category and no points if they are above T_1. We designate PHAs at T_1 and the two percentiles below T_1 as having increased incentives to reduce their ratio (*UTIL-Down* = 1). Given the effects of reversing accruals in future periods, PHAs at the very low end or at the very high end of the distribution have an incentive to manipulate their ratio up. Therefore, we identify PHAs in the first two percentiles and above the 98th percentile as having greater incentives to increase their *UTIL*

---

[24] We also consider PHAs with very high ratios (defined as those with ratios in the highest two percentiles) to have an incentive to increase their ratio with little risk of losing *TRO* points. However, for many PHAs, T_2 is in the highest two percentiles, resulting in overlap in incentive coding. We therefore do not include these high ratio PHAs as having an incentive to increase their ratio.

FIGURE 2
Point Allocation Graph for Tenant Receivables Outstanding



ratio (*UTIL-Up* = 1). All other PHAs are considered to have no incentives related to *UTIL* (*UTIL-Up* = 0 and *UTIL-Down* = 0).

Net Income Fund Balance (NIFB)

Net income as a percent of expendable fund balance (*NIFB*) is a measure of performance. The point allocation, described in Exhibit 2, is dependent on whether the PHA has positive or negative expendable fund balance and an adjusted net income or an adjusted net loss. PHAs with negative expendable fund balance must have positive adjusted net income to earn points. Therefore, PHAs with negative expendable fund balance and small positive adjusted net income, defined as the first two percentiles of positive adjusted net income, are expected to have an incentive to increase their adjusted net income (*NIFB-Up* = 1). PHAs with positive expendable fund balance and adjusted net losses slightly less than or equal to 20 percent of expendable fund balance are identified as having an incentive to reduce their adjusted net loss (*NIFB-Up* = 1). Slightly less than 20 percent is defined as the two percentiles below 20 percent. PHAs with very large positive adjusted net income (above the 98th percentile) and very large adjusted net losses (first two percentiles) are considered to have an incentive to decrease (increase) their net income (loss) and are coded *NIFB-Down* = 1. All other PHAs are considered to have no NIFB incentive (*NIFB-Up* = 0 and *NIFB-Down* = 0).

To summarize, we expect audits to reduce management bias through audit adjustments in the opposite direction of a PHA's incentives to bias information. Because PHAs are unlikely to attempt to manipulate all ratios, or necessarily the same ratios, auditors are unlikely to reduce bias in the same manner across all PHAs. It is an empirical question as to which ratios clients manipulate and which ratios auditors adjust.

Results—Individual Ratio Incentives

Table 4 reports differences in mean audit-related adjustments to ratios between PHAs with and without incentives to manipulate specific ratios. Panel A compares PHAs with an incentive to

EXHIBIT 2

Point Allocation Table for Net Income Expendable Fund Balance

	Negative Expendable Fund Balance	Positive Expendable Fund Balance
Adjusted Net Income	1.5 points	1.5 points
Adjusted Net Loss	0 points	1.5 points if the adjusted net loss is less than or equal to 20 percent of expendable fund balance. 0 points if the adjusted net loss is greater than 20 percent of expendable fund balance.

increase points earned for a ratio to PHAs with no incentive for the corresponding ratio. The results for the current ratio, months expendable fund balance, and net income expendable fund balance are contrary to our expectations. Adjustments to all three ratios are more positive for PHAs with incentive than without incentive. In addition, the difference in audit adjustments to tenant receivables outstanding and utility consumption between PHAs with and without incentive is not significant. Account-level analysis in the next section sheds some light on these unexpected results.

Table 4, Panel B analyzes PHAs that have extreme pre-audit ratio values and have an incentive to attempt to create financial slack in the current year in order to increase the chances of a higher score in future years when accruals reverse. The difference between mean adjustments is significant and in the expected direction for the current ratio and months expendable fund balance, consistent with audits reducing misstatements that create slack in these ratios. The results for tenant receivables outstanding are significant but in the direction opposite to our expectation, while the differences in mean adjustments for utility consumption and net income expendable fund balance are not statistically significant. The results for the current ratio and months expendable fund balance are consistent with the troubled performer incentive tests above. The next section further investigates these as well as the other somewhat mixed results.

VII. ACCOUNT-SPECIFIC ANALYSES

An advantage of our setting is the opportunity to identify specific accounts that are more likely to be used to meet financial ratio targets and thereby receive increased auditor scrutiny under a risk-based audit approach.[25] We identify six such financial statement accounts based on the calculation of FASS ratios and discussions with auditors. As described in Appendix A, the basic accounts used in calculating the FASS ratios are cash, investments, net accounts receivable, tenant receivables, other current assets, current liabilities, tenant revenue, operating revenue, and operating expenses. Our discussions with PHA auditors identified four accounts (cash, investments, receivables from HUD, and revenue from HUD) that are prone to adjustments to correct errors rather than management bias.[26] The adjustments to correct these errors are made by auditors

---

[25] One stream of prior research that has been successful in finding evidence consistent with earnings management examines account and/or transaction specific earnings management (e.g., McNichols and Wilson 1988; Hand 1989; Petroni 1992; Beatty et al. 1995; Jones and Roberts 2006; Keating et al. 2008).

[26] The auditors noted PHAs often forget at year-end to book the receivables and revenue they have earned on cost-reimbursement contracts. The corresponding adjusting entry increases the receivable from HUD and revenue from HUD accounts. In addition, one audit partner noted some investments held for construction-related escrow are classified by PHAs as non-current because they are related to long-term construction even though the investments themselves are theoretically current assets. Consistent with this, we note significant reclassification adjustments between cash and investments that have no effect on the ratio values.

TABLE 4

Comparison of Mean Adjustments in Ratios between PHAs with Incentives to Manipulate Ratios and All Other PHAs

Panel A: Adjustments When Management Has Incentives to *Increase* Points Earned for Ratio[a]

					Mean Adjustment[b]			
Ratio[c]	**Pts.**	**Exp.[d]**	**Incentive n**	**No Incentive n**	**Incentive PHAs**	**No Incentive PHAs**	**Diff.[d]**	**Sig.**
Current Ratio	9	−	2,214	13,232	0.41	−0.32	0.75	###
Months Exp. Fund Bal.	9	−	3,160	12,777	0.13	−0.21	0.34	###
Tenant Rec. Outstand.	4.5	+	9,511	7,526	−0.05	−0.05	−0.01	
Utility Consumption	4.5	+	618	18,176	0.49	0.32	0.16	
Net Inc. Exp. Fund Bal.	1.5	−	505	19,064	0.09	0.02	0.07	###

Panel B: Adjustments When Management Has Incentives to Bias Ratio to *Create Slack* with No Effect on Points Earned[e]

					Mean Adjustment[b]			
Ratio[c]	**Pts.**	**Exp.[d]**	**Incentive n**	**No Incentive n**	**Incentive PHAs**	**No Incentive PHAs**	**Diff.[d]**	**Sig.**
Current Ratio	9	+	385	13,232	0.15	−0.33	0.48	***
Months Exp. Fund Bal.	9	+	514	12,777	1.05	−0.21	1.26	***
Tenant Rec. Outstand.	4.5	−	2,532	7,526	0.23	−0.05	0.27	###
Utility Consumption	4.5	−	775	18,176	0.45	0.32	0.13	
Net Inc. Exp. Fund Bal.	1.5	+	698	18,871	0.02	0.02	0.00	

*** Change for incentive PHAs is significantly different in the expected direction from no-incentive PHAs at p-value < 0.01. Results reported are one-tailed.

Change for incentive PHAs is significantly different but in the unexpected direction from no-incentive PHAs at p-value < 0.01. Results reported are two-tailed.

[a] Panel A compares PHAs with incentives to manipulate these ratios to increase points to those PHAs with no incentive to manipulate the ratios.

[b] t-tests are calculated using the mean change in the ratio.

[c] Ratio formulas are included in Appendix A.

[d] The points represent the maximum possible score for the ratio. The expected sign of the difference is based on our belief that audits reduce manipulation of data through adjusting entries. The differences in the mean changes between incentive PHAs and no-incentive PHAs are calculated from original, unrounded data and are compared to zero.

[e] PHAs with extreme ratio values may have an incentive to manipulate their ratios in a way to create slack in the current year that can be used to increase their FASS score in future years when accruals reverse. Panel B compares PHAs with an incentive to create slack to PHAs with no incentives to manipulate the ratios.

regardless of PHA incentives, and tend to increase the account balances. Since the focus in this section is on whether audits make adjustments that reduce financial bias, we exclude these "error" accounts from our analysis. The six accounts we identify as more likely to be the target of management bias and auditor scrutiny are: tenant receivables (*TENAR*), other current assets (*OCASS*), tenant revenue (*TENREV*), other revenue (*OREV*), current liabilities (*CLIAB*), and total expenses (*EXP*), where other current assets is defined as total current assets less cash, investments, tenant receivables, and HUD receivables, and other revenue is defined as total revenue less tenant revenue and revenue from HUD.[27]

Univariate Analysis

We revisit the results of our ratio tests by examining audit adjustments to the underlying financial statement accounts identified as more likely to be subject to management bias. We compare adjustments of incentive PHAs to no-incentive PHAs and expect greater adjustments for incentive PHAs in the opposite direction of their incentive. We analyze the six identified accounts in a manner consistent with Table 4, and present the results in Table 5.[28]

When PHAs have an incentive to increase the current ratio, audits reduce tenant receivables (*TENAR*), but make adjustments contrary to expectations in current liabilities (*CLIAB*). When PHAs have an incentive to increase months expendable fund balance, audits reduce tenant receivables and increase current liabilities, both consistent with expectations. With incentives to earn more points from tenant receivables outstanding, audits reduce both tenant revenue (consistent with expectations) and tenant receivables (contrary to expectations). This is likely due to the conflicting effects tenant receivables has on the current ratio, months expendable fund balance, and net income expendable fund balance. With incentives to decrease utility consumption, audit adjustments increase expenses (EXP). Finally, in the presence of PHA incentives to increase net income expendable fund balance, audit adjustments increase expenses and current liabilities.

Table 5, Panel B presents the results when PHAs have an incentive to create slack in their ratio values. The adjustment to current liabilities when there are incentives related to the current ratio or months expendable fund balance and the adjustments to other current assets and expenses when there are incentives related to net income expendable fund balance are in the expected direction. However, adjustments to the other variables are either not significant or are significant in the wrong direction. We consider explanations for these mixed results below.

Multivariate Analysis

We use regression analysis to "let the data speak" in terms of evaluating the interplay of troubled performer and individual ratio incentives, non-incentive-based factors affecting adjustments, and the trade-offs PHAs and auditors make in determining which adjustments to record. We model the audit-related adjustment for each of the six financial statement accounts as follows:

$$\Delta DV_t = a_0 + \Sigma_i a_i Incentive\ Variables + \Sigma_c \beta_c Control\ Variables + \varepsilon.$$

Dependent Variables

As discussed above, we identify the following accounts as susceptible to manipulation by management to meet FASS ratio targets: tenant receivables (*TENAR*), other current assets (*OCASS*), revenue from tenants (*TENREV*), other revenue (*OREV*), current liabilities (*CLIAB*), and total expenses (*EXP*). The dependent variables (DV_t) are the audit-related adjustments as measured by

---

[27] These accounts are consistent with those identified in Kreutzfeldt and Wallace (1986), Wright and Ashton (1989), and Houghton and Fogarty (1991) as containing the majority of adjustments.

[28] We perform a similar analysis for the troubled performer incentives presented in Table 3. Most of the adjustments are insignificantly different across the incentive and no-incentive groups (untabulated). Of the 18 comparisons, only two are significantly different in the predicted direction (current liabilities when the incentive was to increase the PHAS score and other revenue when the incentive is to decrease the FASS score) and two are significantly different in the opposite direction (other revenue when the incentive was to increase the PHAS score and expenses when the incentive is to decrease the FASS score).

TABLE 5

Comparison of Mean Adjustments in Variables Included in FASS Ratio Calculations between PHAs with Incentives to Manipulate Ratios and All Other PHAs

Panel A: Adjustments When Management Has Incentives to *Increase* Points Earned for Ratios[a]

				Mean Adjustment[b]			
Ratio/Variable[c]	**Expect.[d]**	**Incentive n**	**No Incentive n**	**Incentive PHAs**	**No Incentive PHAs**	**Diff.[d]**	**Sig.**
Current Ratio							
TENAR	–	2,214	13,232	−187	22	−208	***
OCASS	–	2,214	13,232	−452	−4,769	4,318	
CLIAB	+	2,214	13,232	−8,769	12,326	−21,095	###
Months Expendable Fund Balance							
TENAR	–	3,160	12,777	−96	10	−107	**
OCASS	–	3,160	12,777	−1,681	−4,407	2,726	
CLIAB	+	3,160	12,777	24,824	7,254	17,570	***
EXP	+	3,160	12,777	15,164	16,999	−1,835	
Tenant Receivables Outstanding							
TENAR	+	9,511	7,526	−149	68	−218	###
TENREV	–	9,511	7,526	−660	1,473	−2,133	***
Utility Consumption							
EXP	+	618	18,176	61,291	14,934	46,357	***
Net Income Expendable Fund Balance							
TENAR	–	505	19,064	−166	−11	−155	
OCASS	–	505	19,064	−2,807	−4,975	2,888	
CLIAB	+	505	19,064	28,335	7,370	20,965	**
TENREV	–	505	19,064	1,888	224	1,664	
OREV	–	505	19,064	21,884	12,031	9,852	
EXP	+	505	19,064	29,875	16,899	12,976	*

Panel B: Adjustments When Management Has Incentives to Bias Ratio to *Create Slack* with No Change in Points Earned[e]

				Mean Adjustment[b]			
Ratio/Variable[c]	**Expect.[d]**	**Incentive n**	**No Incentive n**	**Incentive PHAs**	**No Incentive PHAs**	**Diff.[d]**	**Sig.**
Current Ratio							
TENAR	+	385	13,232	169	22	148	
OCASS	+	385	13,232	−35,820	−4,769	−31,051	##
CLIAB	–	385	13,232	−52,910	12,326	−65,235	***
Months Expendable Fund Balance							
TENAR	+	514	12,777	31	10	20	
OCASS	+	514	12,777	−23,623	−4,407	−19,217	
CLIAB	–	514	12,777	−39,557	7,254	−46,811	***
EXP	–	514	12,777	24,374	16,999	7,376	
Tenant Receivables Outstanding							
TENAR	–	2,532	7,626	241	68	173	###
TENREV	+	2,532	7,626	164	1,473	−1,309	##

(continued on next page)

TABLE 5 (continued)

Ratio/Variable[c]	Expect.[d]	Incentive n	No Incentive n	Mean Adjustment[b] Incentive PHAs	No Incentive PHAs	Diff.[d]	Sig.
Utility Consumption							
EXP	–	775	18,176	36,027	14,934	21,092	###
Net Income Expendable Fund Balance							
TENAR	+	698	18,871	–46	–14	–31	
OCASS	+	698	18,871	14,745	–5,627	20,373	**
CLIAB	–	698	18,871	39,645	6,737	32,908	###
TENREV	+	698	18,871	–72	280	–352	
OREV	+	698	18,871	–4,354	12,901	–17,255	##
EXP	–	698	18,871	2,601	17,775	–15,174	**

***, **, * Change for incentive PHAs is significantly different in the expected direction from no-incentive PHAs at p-value < 0.01, 0.05, or 0.10, respectively. Results reported are one-tailed.

###, ## Change for incentive PHAs is significantly different but in the unexpected direction from no-incentive PHAs at p-value < 0.01 and 0.05, respectively. Results reported are two-tailed.

[a] Panel A compares PHAs with incentives to manipulate these ratios to *increase* their FASS to those PHAs with no incentive to manipulate the ratios.

[b] t-tests are calculated using the mean change in the account. t-tests can be greatly affected by extreme values. Therefore, the data were winsorized at the 1st and 99th percentiles.

[c] Variables defined in Appendix B. Ratio formulas are included in Appendix A.

[d] The expected sign of the difference is based on our belief that audits reduce manipulation of data through adjusting entries. The differences in the mean changes between incentive PHAs and no-incentive PHAs are calculated from original, unrounded data and are compared to zero.

[e] PHAs with extreme ratio values may have an incentive to manipulate their ratios in a way to create slack in the current year that can be used to increase their FASS in future years when accruals reverse. Panel B compares PHAs with an incentive to create slack to PHAs with no incentives to manipulate the ratios.

the change from pre- to post-audit values in these accounts, $\Delta TENAR_t$, $\Delta OCASS_t$, $\Delta TENREV_t$, $\Delta OREV_t$, $\Delta CLIAB_t$, and ΔEXP_t.

Management Incentive Variables

We expect PHAs just avoiding the troubled performer classification and just meeting FASS ratio thresholds to have greater incentives to bias financial information. If audits reduce management bias in this setting, then we expect audit-related adjustments to the relevant financial statement accounts. We include indicator variables for the PHAs identified as having greater incentives to bias information. Specifically, we include *P60-Up*, *F-Up*, *F-Down*, *CR-Up*, *CR-Down*, *ME-Up*, *ME-Down*, *TRO-Up*, *TRO-Down*, *UTIL-Up*, *UTIL-Down*, *NIFB-Up*, and *NIFB-Down*, all defined above and summarized in Appendix B. The expected sign of the association between audit adjustments and client incentives depends on the financial statement account and how it is used in achieving financial targets.[29] For example, tenant receivables (*TENAR*) is used in the calculation of all ratios except utility consumption. PHAs with an incentive to increase their points prefer higher values of *TENAR* for the current ratio, months expendable fund balance, and net income expendable fund balance, and lower values for tenant receivables outstanding. If audits reduce bias to meet these incentives, then we should observe a negative association between $\Delta TENAR$ and *CR-Up*, *ME-Up*, and/or *NIFB-Up* and a positive association between $\Delta TENAR$ and

---

[29] Details of the ratio calculations are provided in Appendix A.

TRO-Down. The reverse holds when the incentives are in the opposite direction. When the incentive is to increase or decrease overall points earned (*P60-Up*, *F-Up*, or *F-Down*) there is no clear prediction as to how the audit would reduce management bias in *TENAR* due to these conflicting effects, and therefore we do not predict signs. Predictions for the other dependent variables follow the same logic, but lack the conflicting issues characterizing *TENAR*; expectations for all variables are indicated in Table 6. We expect a statistically significant association between audit-related adjustments and some but not necessarily all of the conjectured incentives to bias financial information.[30]

Control Variables

The magnitude of audit adjustments is likely also affected by several factors unrelated to management bias. We expect a greater scope for audit-related changes in large and complex entities without any specific intent of management to bias the financial data. Consistent with this expectation, Kreutzfeldt and Wallace (1986) and Hylas and Ashton (1982) find that the magnitude of errors as a percent of total assets varies by size of the client. In addition, Kinney and McDaniel (1989) find firms with restatements of quarterly data are smaller than non-restatement firms within their industry. We include *LNA*, the natural log of assets at the end of the year, and *#PROG*, the total number of federal programs the PHA manages, to control for the size and complexity of the PHA in all six equations. We also include *UNITS*, the number of apartment units times the number of months each unit was available as an additional measure of size in four of the six equations where the underlying account may be proportional to the number of apartments (e.g., tenant receivables and expenses but not other current assets).

Weak internal controls, poor accounting, and financial instability could affect the level of and resulting adjustment to financial statement accounts without any specific management intent to bias the financial data. Kreutzfeldt and Wallace (1986) find that the incidence of errors is higher for firms with poor controls. Further, Kinney and McDaniel (1989) find that firms restating quarterly data are more likely to receive a going-concern opinion. We include measures of internal control quality, IC_{t-1}, and other financial reporting issues, $PROB_{t-1}$. IC_{t-1} is equal to 1 if the PHA had reportable conditions or material weaknesses reported in the previous year, 0 otherwise, and $PROB_{t-1}$ is 1 if the PHA received a going-concern audit opinion, an other-than-unqualified audit opinion, indications of findings or questioned costs, or indications of material noncompliance with federal regulations in the prior year, 0 otherwise.[31] We include both measures in all six equations.

We also control for the likely effectiveness of the auditor in identifying and correcting misstatements. DeAngelo (1981) hypothesizes that audit quality improves with auditor size, and significant prior research finds that corporations and users of financial statements believe large auditors provide higher quality audits (e.g., Craswell et al. 1995; Beattie et al. 2001; Allen 1994; Balvers et al. 1988; Beatty 1989; Teoh and Wong 1993). While most prior research uses the Big 5/non-Big 5 designation for measures of audit quality, we use Top 50/non-Top 50 because of the significant diversity in firms auditing PHAs. *TOP50* is 1 if the auditor is one of the top 50 auditors as ranked by *Accounting Today*, 0 otherwise.[32] We include the natural log of the audit fee (*FEE*) as an additional control for auditor effectiveness (Copley et al. 1994; Abbott et al. 2006). Finally, we

---

[30] For all the associations to be significant, PHAs would have to bias all six accounts in the same direction for each incentive. Further, auditors would need to adjust all six accounts in the same direction for each incentive. Given the degrees of freedom available to PHA managers and auditors, it is unlikely that all will respond in exactly the same manner.

[31] Other-than-unqualified opinions include qualified, adverse, and disclaimers of opinion. It does not include unqualified opinions modified for consistency, emphasis of a matter, or going-concern issues.

[32] We also estimate the regressions using the Big 5 designation with no change in the results.

TABLE 6

Regressions Run Using OLS[a]

Ind. Var.[c]		ΔTENAR[b]			ΔOCASS[b]			ΔTENREV[b]			ΔOREV[b]			ΔCLIAB[b]			ΔEXP[b]	
Incentives Related to Increasing Overall Score																		
P60-Up	?	−278		–	−33		–	−1,405		–	24,718		+	42,989		+	14,465	
F-Up	?	194		–	−41,948	**	–	135		–	−21,016	*	+	−33,058		+	4,110	
CR-Up	–	−246	*	–	8,314								+	−50,093	###			
ME-Up	–	−24		–	783								+	32,450	***	+	−147	
TRO-Down	+	−339	###				–	−1,183	**									
UTIL-Down																+	35,698	**
NIFB-Up	–	−87		–	21,482		–	2,123		–	16,019		+	39,000	**	+	12,167	
Incentives Related to Creating Slack																		
F-Down	?	801		+	−2,917		+	−1,452		+	18,682		–	−39,514		–	53,560	##
CR-Down	+	−140		+	−2,072								–	529				
ME-Down	–	79		+	−3,323								–	−12,429	*	–	1,041	
TRO-Up	–	201	##				+	111										
UTIL-Up																–	18,366	
NIFB-Down	+	−23		+	33,318	*	+	−1,725		+	−37,354	##	–	49,940	###	–	−39,861	***
Controls[c]																		
DV_{t-1}[d]		0.11	***		0.01			0.21	***		−0.01			0.12	***		0.01	***
UNITS		0.00	**					0.00			0.24	***					0.11	
LNA		59			−656			45			7,400	**		12,079	**		8,528	***
#PROG		−19			−1,335			−121			3,762	**		−1,003			3,653	**
IC_{t-1}		54			3,410			−219			2,193			9,511			18,652	**
$PROB_{t-1}$		−2			−17,712	**		479			2,819			−560			6,515	
TOP50		−347			−31,671			2,171			15,811			41,489			−14,440	
FEE		−41			−800			480	**		−9,196	***		1,825			3,066	
Constant		−274			21,912			−3,810			−44,868			−190,021	***		−165,592	***
R^2		1.85%			0.38%			5.09%			1.96%			2.63%			3.60%	
n = 13,009																		

***, **, * Coefficient estimate is significantly different from zero at p-value < 0.01, 0.05, and 0.10, respectively. Tests are one-tailed if coefficient sign is consistent with expectations, two-tailed if there is no expectation.

(continued on next page)

TABLE 6 (continued)

###, ## Coefficient estimate is significantly different from zero with a sign *not* consistent with expectations at p-value < 0.01 and 0.05, respectively. Results reported are two-tailed.

[a] The following six regressions were run simultaneously using ordinary least squares estimation, controlling for the lack of independence across observations using a clustering approach:

$$\Delta TENAR = \alpha_0 + \alpha_1 P60\text{-}Up + \alpha_2 F\text{-}Up + \alpha_3 F\text{-}Down + \alpha_4 CR\text{-}Up + \alpha_5 CR\text{-}Down + \alpha_6 ME\text{-}Up + \alpha_7 ME\text{-}Down + \alpha_8 TRO\text{-}Up + \alpha_9 TRO\text{-}Down + \alpha_{10} \Delta TENAR_{t-1} + \alpha_{11} UNITS + \alpha_{12} LNA + \alpha_{13} \#PROG + \alpha_{14} IC_{t-1} + \alpha_{15} PROB_{t-1} + \alpha_{16} TOP50 + \alpha_{17} FEE + \varepsilon_1;$$

$$\Delta OCASS = \gamma_0 + \gamma_1 P60\text{-}Up + \gamma_2 F\text{-}Up + \gamma_3 F\text{-}Down + \gamma_4 CR\text{-}Up + \gamma_5 CR\text{-}Down + \gamma_6 ME\text{-}Up + \gamma_7 ME\text{-}Down + \gamma_8 \Delta OCASS_{t-1} + \gamma_9 LNA + \gamma_{10} \#PROG + \gamma_{11} IC_{t-1} + \gamma_{12} PROB_{t-1} + \gamma_{13} TOP50 + \gamma_{14} FEE + \varepsilon_3;$$

$$\Delta TENREV = \delta_0 + \delta_1 P60\text{-}Up + \delta_2 F\text{-}Up + \delta_3 F\text{-}Down + \delta_4 TRO\text{-}Up + \delta_5 TRO\text{-}Down + \delta_6 NIFB\text{-}Up + \delta_7 NIFB\text{-}Down + \delta_8 \Delta TENREV_{t-1} + \delta_9 UNITS + \delta_{10} LNA + \delta_{11} \#PROG + \delta_{12} IC_{t-1} + \delta_{13} PROB_{t-1} + \delta_{14} TOP50 + \delta_{15} FEE + \varepsilon_4;$$

$$\Delta OREV = \theta_0 + \theta_1 P60\text{-}Up + \theta_2 F\text{-}Up + \theta_3 F\text{-}Down + \theta_4 NIFB\text{-}Up + \theta_5 NIFB\text{-}Down + \theta_6 \Delta OREV_{t-1} + \theta_7 UNITS + \theta_8 LNA + \theta_9 \#PROG + \theta_{10} IC_{t-1} + \theta_{11} PROB_{t-1} + \theta_{12} TOP50 + \theta_{13} FEE + \varepsilon_5;$$

$$\Delta CLIAB = \varphi_0 + \varphi_1 P60\text{-}Up + \varphi_2 F\text{-}Up + \varphi_3 F\text{-}Down + \varphi_4 CR\text{-}Up + \varphi_5 CR\text{-}Down + \varphi_6 ME\text{-}Up + \varphi_7 ME\text{-}Down + \varphi_8 \Delta CLIAB_{t-1} + \varphi_9 LNA + \varphi_{10} \#PROG + \varphi_{11} IC_{t-1} + \varphi_{12} PROB_{t-1} + \varphi_{13} TOP50 + \varphi_{14} FEE + \varepsilon_6;$$

$$\Delta EXP = \lambda_0 + \lambda_1 P60\text{-}Up + \lambda_2 F\text{-}Up + \lambda_3 F\text{-}Down + \lambda_4 ME\text{-}Up + \lambda_5 ME\text{-}Down + \lambda_6 UTIL\text{-}Up + \lambda_7 UTIL\text{-}Down + \lambda_8 NIFB\text{-}Up + \lambda_9 NIFB\text{-}Down + \lambda_{10} \Delta EXP_{t-1} + \lambda_{11} UNITS + \lambda_{12} LNA + \lambda_{13} \#PROG + \lambda_{14} IC_{t-1} + \lambda_{15} PROB_{t-1} + \lambda_{16} TOP50 + \lambda_{17} FEE + \varepsilon.$$

[b] Dependent variables are the change from pre- to post-audit value calculated as audited value – unaudited value.

[c] Variable definitions in Appendix B.

[d] DV_{t-1} is defined as the prior year value of the dependent variable for each regression.

include the lagged value of the dependent variable, as audit-related adjustments in the current year are likely to be similar to adjustments in the prior year (Kinney 1979; Hylas and Ashton 1982; Kreutzfeldt and Wallace 1986; Houghton and Fogarty 1991).

In summary, we model the audit-related adjustments to six financial statement accounts as a function of management incentives to bias financial reporting and non-incentive-based factors affecting adjustments. Since there is a lack of independence across observations, given each PHA is included in the dataset multiple times, we estimate the equations using a clustering approach in OLS that adjusts standard errors for this lack of independence (STATA 2007).[33]

Results—Multivariate Analysis

Table 6 presents the results of estimating the six audit adjustment equations. We analyze the table across the rows to identify which accounts auditors constrain in light of management's incentives. Consistent with prior tables, we group the incentives by "increasing scores" and "creating financial slack." None of the identified accounts appear to be systematically targeted for adjustment when management has incentives to increase the overall PHAS score (*P60-Up*), while audits seem to reduce bias in other current assets (*OCASS*) and other revenue (*OREV*) for incentives to increase the FASS score (*F-Up*). These results are in contrast to the ratio-level results presented in Table 3, Panels A and B, which suggest a broader auditor focus on potential manipulation in the current ratio and months expendable fund balance.

Given the significant overlap in variables used in calculating the current ratio and months expendable fund balance and our previous results, we expect auditors to reduce bias in a similar manner for each of the ratios. However, this is not the case. Auditors appear to focus on tenant receivables (*TENAR*) when management has an incentive to increase the current ratio (*CR-Up*), but focus on current liabilities (*CLIAB*) when management has an incentive to increase months expendable fund balance (*ME-Up*). Adjustments to *CLIAB* are contrary to expectation when the incentive is to increase the current ratio.

Audits focus on adjusting tenant revenue (*TENREV*) rather than tenant receivables (*TENAR*) when management has an incentive to increase points earned from tenant receivables outstanding (*TRO-Down*). This is not surprising given that increases in *TENAR* will increase the current ratio and months expendable fund balance, both of which are worth 9 points compared to 4.5 for *TRO*. Further, audits appear to reduce potential bias in total expenses when management has an incentive to increase points earned for utility consumption (*UTIL-Down*), and they appear to reduce potential bias in current liabilities when management's incentive is to increase the net income expendable fund balance points (*NIFB-Up*). Consistent with the account-level results in Table 5, Panel A, it appears that some consideration is given to a PHA's incentive to misreport in determining the adjustment to each of the identified accounts. This contrasts with the general lack of results in Table 4, Panels A and B. Below we discuss possible explanations.

Finally, consistent with results from Table 5, Panel B (but in contrast to Table 3, Panel C and Table 4, Panel B), audits do not appear to reduce potential manipulation related to creating financial slack as seen by the mostly nonsignificant results or significant results in the opposite direction expected for all of the incentives. The exception is when management has increased incentives to manage net income expendable fund balance down (*NIFB-Down*). In this situation, auditors increase other current assets (*OCASS*) and reduce expenses (*EXP*).[34]

---

[33] A secondary issue is the potential correlation in error terms across the equations due to the nature of audit adjustments. One way to address this is to use a seemingly unrelated regression (SUR). SUR extends ordinary least squares analysis to estimate a system of linear equations with correlated error terms increasing the efficiency of the parameter estimates. Unfortunately, we are unable to implement SUR and control for the lack of independence in panel data simultaneously, so we estimate the equations using both methods. The results across OLS and SUR are substantively similar.

Overall, our results support the H2 prediction that audits reduce management bias when there are greater incentives to meet financial targets. Auditors appear to focus more on potential manipulation that would increase points earned rather than potential manipulation that creates slack. This asymmetric response is consistent with prior research showing auditors react more strongly to income-increasing rather than income-decreasing earnings management (Abbott et al. 2006).

Discussion

We identify two possible reasons (not mutually exclusive) for the conflicting results between account- and ratio-level analyses. First, during audit fieldwork, auditors are likely to focus on the effects of incentives on individual accounts and ratios. The effects of the PHAS and FASS score incentives are likely considered during the completion of the audit when evaluating the overall presentation of the financial statements. This higher level analysis at the end of the audit might result in changes to ratios as a whole, but not in any one account systematically. Second, while the ratio-level results for incentives related to creating slack presented in Tables 3 and 4 are consistent with our hypotheses, they are also consistent with auditors making error-correcting, account-increasing adjustments to the four accounts (cash, investments, HUD receivables, and HUD revenue) noted in discussions with PHA auditors. When we focus on adjustments to individual accounts most likely to contain management bias, the results on incentives to create slack are eliminated. We interpret the conflicting results between the account- and ratio-level analyses as confirming the importance of understanding the industry. Finally in addition to the factors already discussed, the mixed results within analyses could be an indication that auditors are not aware of all incentives, audits are unable to reduce some management manipulation, all PHAs do not choose the same accounts to manipulate, and/or all auditors do not choose the same accounts to adjust in order to reduce manipulation.

VIII. SENSITIVITY ANALYSIS

In measuring client incentives and audit-related adjustments, our analyses reflect a number of decisions that could potentially affect the results. We evaluate the sensitivity of our results to three of these choices here. First, we repeat the tests documented in Tables 3 through 6 using a one-percentile window above/below the incentive thresholds rather than the two-percentile window used in the reported results. Next, we repeat the tests in Tables 5 and 6 using audit adjustments scaled by total assets and total revenues. Finally, we repeat the tests in Tables 4 and 5 after expanding the incentive group to include PHAs that just missed HUD thresholds for increased points. While the specific results on the individual variables are not consistent across all measurement methods and although there are a few results in contrast to expectations (consistent with the tabulated results), the overall tenor of the conclusions for each test is unchanged. Audits appear to make corrections to client accounts in a manner consistent with reducing management bias with a focus on overstatement of scores.[35]

In addition, we analyze two subsets of our sample population. First, we perform the analyses for Tables 3 through 6 using only those observations from September 2001 through June 2003, as our data on PHAS and FASS scores are most complete over this time period and the incentives related to current ratio and months expendable fund balance are strongest (9 points if the ratio is 1.0 or above, otherwise 0 points). Second, we identify all observations for which client incentives could be considered in conflict with one another. For example, it is possible for a PHA to have an incentive to increase the current ratio at the same time they have an incentive to decrease months expendable fund balance. Because many of the same accounts are used in both ratios, it is difficult to determine which incentive would outweigh the other. While we feel this is a contribution of our multivariate tests, we repeat the analyses using only observations with no conflicting incentives. The results from both of these additional analyses are generally consistent with the results presented.

---

[34] Tests of multicollinearity for all variables in Table 6 reveal the highest variance inflation factor is 2.79 for *LNASSET* in the *TENAR* regression, which is well below 10.0, the level designated in Kennedy (1997) as cause for concern.

[35] Results available from authors upon request.

Finally, we perform the tests for Tables 3 through 5 year by year to control for the lack of independence in the univariate tests as a result of individual PHAs being in the sample multiple times. Although individual results each year are somewhat different from those reported, taken together the overall results are consistent with the conclusions previously discussed.

IX. CONCLUSION

Audits are widely believed to be a means of improving financial reporting and mitigating agency costs. However, it is often difficult to determine the effect of an audit because only the end results, the audited financial statements and audit report, are generally publicly available. We take advantage of the unique reporting requirements of public housing authorities to examine whether audits result in significant adjustments to the financial statements and whether those adjustments reflect auditor sensitivity to management incentives to bias the financial statements.

We find that audits result in economically and statistically significant increasing, as well as decreasing, adjustments across the financial statements. Audits matter in this nonprofit, low litigation risk setting. This is consistent with prior research on audit adjustments by large audit firms across various for-profit industries. Further, we find an overall pattern of audit adjustments consistent with auditors reducing potential management bias in the financial statements. We also find patterns consistent with auditors attending more to potential overstatement than understatement of scores, consistent with the pattern of asymmetric audit reactions found in prior literature. Finally, methodologically, our conflicting results using aggregated (ratio-level) and disaggregated (account-level) financial data provide support for the importance of expanding audit research beyond financial statement results. Our results should be of interest to academics, regulators, and standard-setters as part of a growing body of research on financial misreporting in nonprofit settings as well as the role of auditors as financial intermediaries.

Future research could investigate the effects of the length of time between reports on differences in the pre- and post-audit data. While we believe the audit adjustments are not a result of subsequent events, it is possible some adjustments are a result of the long time period between pre- and post-audit filings.[36] Given the diverse nature of the audit market for nonprofit organizations, measures of quality other than size might be more appropriate and might provide additional insight on which auditors better constrain manipulation. Finally, future research might investigate the effects of client incentives and potential manipulation on audit fees.

REFERENCES

Abbott, L. J., S. Parker, and G. F. Peters. 2006. Earnings management, litigation risk, and asymmetric audit fee responses. *Auditing: A Journal of Practice & Theory* 25 (May): 85–98.

Allen, A. C. 1994. The effect of large-firm audits on municipal bond rating decisions. *Auditing: A Journal of Practice & Theory* 13 (Spring): 115–125.

---

[36] We thank an anonymous reviewer for this suggestion.

American Institute of Certified Public Accountants (AICPA). 1972. *Responsibilities and Functions of the Independent Auditor*. Statement on Auditing Standard No. 1. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2002. *Consideration of Fraud in a Financial Statement Audit*. Statement on Auditing Standard No. 99. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2006a. *Understanding the Entity and Its Environment and Assessing the Risks of Material Misstatement*. Statement on Auditing Standard No. 109. New York, NY: AICPA.

American Institute of Certified Public Accountants (AICPA). 2006b. *Performing Audit Procedures in Response to Assessed Risks and Evaluating the Audit Evidence Obtained*. Statement on Auditing Standard No. 110. New York, NY: AICPA.

Balsam, S., J. Krishnan, and J. K. S. Yang. 2003. Auditor industry specialization and earnings quality. *Auditing: A Journal of Practice & Theory* 22 (September): 71–97.

Balvers, R. J., B. McDonald, and R. E. Miller. 1988. Underpricing of new issues and the choice of auditor as a signal of investment banker reputation. *The Accounting Review* 63 (October): 605–622.

Beattie, V., A. Goodacre, K. Pratt, and J. Stevenson. 2001. The determinants of audit fees—Evidence from the voluntary sector. *Accounting and Business Research* 31 (Autumn): 243–274.

Beatty, A., S. L. Chamberlain, and J. Magliolo. 1995. Managing financial reports of commercial banks: The influence of taxes, regulatory capital, and earnings. *Journal of Accounting Research* 33 (Autumn): 231–261.

Beatty, R. P. 1989. Auditor reputation and the pricing of initial public offerings. *The Accounting Review* 64 (October): 693–709.

Becker, C. L., M. L. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (Spring): 1–24.

Bell, T., and W. Knechel. 1994. Empirical analyses of errors discovered in audits of property and casualty insurers. *Auditing: A Journal of Practice & Theory* 13 (1): 84–100.

Bell, T. B., W. R. Knechel, and J. J. Willingham. 1992. An empirical investigation of the relationship between selected attributes of accounting systems and the incidence of audit differences. Working paper, KPMG Peat Marwick.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (December): 99–126.

Craswell, A. T., J. R. Francis, and S. L. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20 (December): 297–322.

Chung, R., M. Firth, and J. B. Kim. 2005. Earnings management, surplus free cash flow, and external monitoring. *Journal of Business Research* 58 (6): 766–776.

Collins, J. H., D. S. Shackelford, and J. M. Wahlen. 1995. Bank differences in the coordination of regulatory capital, earnings, and taxes. *Journal of Accounting Research* 30 (Autumn): 263–291.

Copley, P. A., M. S. Doucet, and K. M. Gaver. 1994. A simultaneous equations analysis of quality control review outcomes and engagement fees for audits of recipients of federal financial assistance. *The Accounting Review* 69 (January): 244–256.

DeAngelo, L. E. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (December): 183–199.

DeFond, M. L., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17 (January): 145–176.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *The Journal of Business* 72 (1): 1–33.

Duggan, M. G. 2000. Hospital ownership and public medical spending. *The Quarterly Journal of Economics* 115 (November): 1343–1373.

Hand, J. 1989. Did firms undertake debt-equity swaps for an accounting paper profit or true financial gain? *The Accounting Review* 64 (October): 587–623.

Healy, P. M. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (April): 85–107.

Healy, P. M., and J. M. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (December): 365–383.

Hirst, D. E. 1994. Auditor sensitivity to earnings management. *Contemporary Accounting Research* 11 (Fall): 405–422.

Houghton, C. W., and J. A. Fogarty. 1991. Inherent risk. *Auditing: A Journal of Practice & Theory* 10 (Spring): 1–21.

Hylas, R. E., and R. H. Ashton. 1982. Audit detection of financial statement errors. *The Accounting Review* 57 (October): 751–765.

Iceman, R. C., and W. A. Hillison. 1991. Disposition of audit-detected errors: Some evidence on evaluative materiality. *Auditing: A Journal of Practice & Theory* 10 (Spring): 22–34.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs, and ownership structure. *Journal of Financial Economics* 3 (October): 305–360.

Jones, C. L., and A. A. Roberts. 2006. Management of financial information in charitable organizations: The case of joint-cost allocations. *The Accounting Review* 81 (January): 159–178.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Keating, E. K., L. M. Parsons, and A. A. Roberts. 2008. Misreporting fundraising: How do nonprofit organizations account for telemarketing campaigns? *The Accounting Review* 83 (March): 417–446.

Kennedy, P. 1997. *A Guide to Econometrics. 3rd edition.* Cambridge, MA: The MIT Press.

Kim, J.-B., R. Chung, and M. Firth. 2003. Auditor conservatism, asymmetric monitoring, and earnings management. *Contemporary Accounting Research* 20 (Summer): 323–359.

Kinney, W. R., Jr. 1979. The predictive power of limited information in preliminary analytical review: An empirical study. *Journal of Accounting Research* 17 (Supplement): 148–165.

Kinney, W. R., Jr., and R. D. Martin. 1994. Does auditing reduce bias in financial reporting? A review of audit-related adjustments studies. *Auditing: A Journal of Practice & Theory* 13 (Spring): 149–156.

Kinney, W. R., Jr., and L. S. McDaniel. 1989. Characteristics of firms correcting previously reported quarterly earnings. *Journal of Accounting and Economics* 11 (February): 71–93.

Kreutzfeldt, R. W., and W. A. Wallace. 1986. Error characteristics in audit populations: Their profile and relationship to environmental factors. *Auditing: A Journal of Practice & Theory* 6 (Fall): 20–43.

Krishnan, R., M. H. Yetman, and R. J. Yetman. 2006. Expense misreporting in nonprofit organizations. *The Accounting Review* 81 (March): 399–420.

Leone, A. J., and R. L. Van Horn. 2005. How do non-profit hospitals manage earnings? *Journal of Health Economics* 24 (4): 815–837.

Libby R., and W. R. Kinney, 2000. Does mandated audit communication reduce opportunistic corrections to manage earnings forecasts? *The Accounting Review* 75 (4): 383–404.

McNichols, M., and G. P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

Messier, W. F., Jr., S. M. Glover, and D. F. Prawitt. 2006. *Auditing and Assurance Services: A Systematic Approach. 5th edition.* New York, NY: McGraw-Hill Irwin.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Ng, T. B. 2007. Auditors' decisions on audit differences that affect significant earnings thresholds. *Auditing: a Journal of Practice & Theory* 26 (1): 71–89.

Office of Management and Budget. 1997. *Audits of States, Local Governments, and Non-Profit Organizations.* Circular No. A-133: Washington, D.C.: The Executive Office of the President.

Petroni, K. R. 1992. Optimistic reporting in the property-casualty insurance industry. *Journal of Accounting and Economics* 15: 485–508.

Scholes, M. S., G. P. Wilson, and M. A. Wolfson. 1990. Tax planning, regulatory capital planning, and financial reporting strategy for commercial banks. *Review of Financial Studies* 3 (4): 625–650.

Sweeney, A. P. 1994. Debt-covenant violations and managers' accounting responses. *Journal of Accounting and Economics* 17 (3): 281–308.

STATA. 2007. *Stata Base Reference Manual, Volume 3 Q-Z, Release 10.* College Station, TX: StataCorp LP.

Teoh, S. H., and T. J. Wong. 1993. Perceived auditor quality and the earnings response coefficient. *The Accounting Review* 68 (April): 346–366.

Watts, R., and J. Zimmerman. 1983. Agency problems, auditing and the theory of the firm: Some evidence. *Journal of Law and Economics* 26 (October) 613–633.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory.* Englewood Cliffs, NJ: Prentice Hall.

Wright A., and R. H. Ashton. 1989. Identifying audit adjustments with attention-directing procedures. *The Accounting Review* 64 (October): 710–728.

Wright A., and S. Wright. 1997. An examination of factors affecting the decision to waive audit adjustments. *Journal of Accounting, Auditing and Finance* 12 (1): 15–36.

Yetman, R. J. 2001. Tax-motivated expense allocations by nonprofit organizations. *The Accounting Review* 76 (July): 297–311.

APPENDIX A

Details of Financial Score Component Ratio Calculations

Ratio	Numerator	Denominator
Current Ratio (*CR*)	Cash + Net Accounts Receivable + Unrestricted Investments + Investments Restricted for Payment of Current Liabilities + Prepaid and Other Assets	Total Current Liabilities − Interprogram Due To
Months Expendable Fund Balance (*ME*)	Cash + Net Accounts Receivable + Unrestricted Investments + Investments Restricted for Payment of Current Liabilities + Prepaid and Other Assets − Total Current Liabilities + Current Portion of Long-Term Debt for Capital Projects + Interprogram Due To − Long-Term Debt, Net of Current for Operating Borrowings	(Debt Principal Payments for Enterprise + Total Operating Expenses + Extraordinary Maintenance + Non-Capitalized Casualty Losses + Debt Principal Payments for Governmental Funds + Dwelling Units Rent Expense[a])/12
Tenant Receivables Outstanding (*TRO*)	Tenant Receivables	Total Tenant Revenue[a]/365
Occupancy Loss (*OL*)	Number of Unit Months Leased	Gross Number of Units[b]
Expense Management/ Utility Consumption (*UTIL*)	0.34 ∗ Administrative Expenses + 0.10 ∗ Tenant Services Expenses + 0.03 ∗ Utilities Expenses + 0.10 ∗ Ordinary Maintenance and Operations + 0.10 ∗ Protective Services Expenses + 0.33 ∗ General Expenses	Number of Unit Months Leased
Net Income Divided by Expendable Fund Balance (*NIFB*)	Excess of Operating Revenue Over Operating Expenses − Capital Grants − Extraordinary Maintenance − Non-Capitalized Casualty Losses − Housing Assistance Payments − Fraud Losses − Capital Outlays for Governmental − Dwelling Units Rent Expense[a]	Cash + Net Accounts Receivable + Unrestricted Investments + Investments Restricted for Payment of Current Liabilities + Prepaid and Other Assets − Total Current Liabilities + Current Portion of Long-Term Debt for Capital Projects + Interprogram Due To − Long-Term Debt, Net of Current for Operating Borrowings

---

[a] Equation provides for adjustments to convert the PHA's ratio to the full accrual basis if the PHA uses the modified accrual basis of accounting.

[b] This fraction is subtracted from 1 to obtain the loss ratio.

APPENDIX B

Definition of Variables

CLIAB	Total current liabilities at end of year
CR	Current ratio; calculation detailed in Appendix A
CR-Up	Coded as 1 if PHA has incentive to increase *CR*; 0 otherwise
CR-Down	Coded as 1 if PHA has incentive to decrease *CR*; 0 otherwise
EXP	Total expenses at end of year
FEE	Natural log of audit fee
F-Up	Coded as 1 if PHA has incentive to increase FASS score; 0 otherwise
F-Down	Coded as 1 if PHA has incentive to decrease FASS score; 0 otherwise
IC	Coded as 1 if auditor identified reportable conditions or material weaknesses in internal controls; 0 otherwise
LNA	Natural log of assets at end of year
ME	Months expendable fund balance; calculation detailed in Appendix A
ME-Up	Coded as 1 if PHA has incentive to increase *ME*; 0 otherwise
ME-Down	Coded as 1 if PHA has incentive to decrease *ME*; 0 otherwise
NIFB	Net income expendable fund balance; calculation detailed in Appendix A
NIFB-Up	Coded as 1 if PHA has incentive to increase *NIFB*; 0 otherwise
NIFB-Down	Coded as 1 if PHA has incentive to decrease *NIFB*; 0 otherwise
OCASS	Total current assets less cash, investments, tenant receivables, and HUD receivables at end of year
OREV	Total revenue less revenue from tenants and revenue from HUD at end of year
P60-Up	Coded as 1 if PHA has incentive to increase PHAS score above 60; 0 otherwise
PROB	Coded as 1 if PHA received a going-concern opinion, other-than-unqualified opinion on the financial statements, indications of findings, indications of questioned costs, or indications of material noncompliance with federal regulations; 0 otherwise
#PROG	Number of federal programs managed by PHA
TENAR	Total accounts receivable from PHA tenants at end of year
TENREV	Total revenue from PHA tenants at end of year
TOP50	Coded as 1 if PHA auditor was in top 50 audit firms as reported by *Accounting Today*; 0 otherwise
TRO	Tenant receivable outstanding; calculation detailed in Appendix A
TRO-Up	Coded as 1 if PHA has incentive to increase *TRO*; 0 otherwise
TRO-Down	Coded as 1 if PHA has incentive to decrease *TRO*; 0 otherwise
UNITS	Number of units available times months available
UTIL	Expense management/utility consumption; calculation detailed in Appendix A
UTIL-Up	Coded as 1 if PHA has incentive to increase *UTIL*; 0 otherwise
UTIL-Down	Coded as 1 if PHA has incentive to decrease *UTIL*; 0 otherwise

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1321–1347

American Accounting Association
DOI: 10.2308/accr-10042

Relative Firm Profitability and Stock Return Sensitivity to Industry-Level News

Shengquan Hao
Shanghai Jiaotong University

Qinglu Jin
Shanghai University of Finance and Economics

Guochang Zhang
Hong Kong University of Science and Technology

ABSTRACT: This study provides theory and evidence to demonstrate how relative firm profitability within an industry affects stock return sensitivity to industry-level news. Extending the Cournot and Bertrand competition models, we predict that (1) the returns of less profitable firms in an industry are more sensitive to industry-level news than those of more profitable firms, and that (2) this inverse relation between relative profitability and return sensitivity is more pronounced when there is positive rather than negative industry news, especially in industries with high (versus low) capital intensity. Using industry returns to proxy for industry-level news, we obtain empirical results consistent with these predictions. We further find that the two fundamental factors that contribute to profitability—cost efficiency and market share—each exhibit an effect similar to that of relative profitability in affecting return sensitivity. Our results remain unchanged after controlling for stock price movements attributable to common risk factors and firm-specific accounting information, and they hold over a range of robustness tests.

Keywords: *relative profitability; industry-level news; return sensitivity; industry beta.*

Data Availability: *The data used in this study are available from public sources.*

We thank two anonymous referees, Peter Chen, Mingyi Hung, Clive Lennox, Paul Zarowin (editor), Xiao-Jun Zhang, and workshop participants at the Chinese University of Hong Kong, Fudan University, Hong Kong University of Science and Technology, University of British Columbia, University of California, Berkeley, and University of Southern California for their helpful comments. Professor Jin acknowledges the financial support of the MOE through the Institute of Finance and Accounting at SUFE (project number 00JJD630006) and of the National Natural Science Foundation of China (project number 71072036).

Editor's note: Accepted by Paul Zarowin.

Submitted: February 2009
Accepted: January 2011
Published Online: April 2011

I. INTRODUCTION

An important challenge for financial analysts and investors is ascertaining how the stock prices of individual firms move with the changing conditions of the sectors in which they operate. Are the stock returns of certain firms within an industry more sensitive to industry-wide news than those of others? If so, then what firm-specific attributes help to predict or explain cross-sectional differences in return sensitivity to industry news? These questions are largely unexplored in the extant literature. In this study, we investigate the usefulness of accounting data in determining a firm's return sensitivity to industry-level news. Although an extensive literature has examined the importance of accounting information *by itself* in explaining stock returns, few studies have considered whether and how (firm-specific) accounting information is also related to the way in which stock prices respond to aggregate-level (such as industry-wide) information, an issue that concerns the interactive effects of the two types of information in valuations.

Understanding such interactive effects is useful not only for predicting or explaining (conditional) firm returns, but also for uncovering the link between accounting information and systematic risk. Previous studies (e.g., Beaver et al. 1970; Rosenberg and McKibben 1973; Beaver and Manegold 1975) have found that systematic market risk (market beta) is correlated with the earnings beta and earnings volatility (see Ryan [1997] for a survey). As Ryan (1997) points out, however, this line of research relies primarily on economic intuition, rather than rigorous theory, to motivate empirical analysis. In this study, we also aim to extend this literature by providing a theoretical explanation for how accounting information is related to one specific form of systematic risk (i.e., return sensitivity to industry news).

This study adopts an industrial organization (IO) perspective to explore the interaction between industry-level and firm-specific information in explaining returns. In practice, firms operating in the same industry behave strategically in competition with one another, thus creating inter-firm dependencies in operational decisions. If a firm's cash flows are generated from decisions that are dependent on the decisions of competing firms, then there should also be inter-firm stock return dependencies that reflect the underlying cash flows. Following this line of reasoning, a firm's competitiveness within a given industry should have consequences for stock return behavior. We focus on relative profitability as an indicator of competitiveness to examine its influence on returns within the industry context.

We first extend the standard Cournot and Bertrand models to examine how firms in the same industry adjust their operations in response to common external shocks, thus resulting in changes in firm value. Our analysis leads to the prediction that the stock returns of less profitable firms are more sensitive to industry-wide shocks than are those of more profitable firms. In addition, this inverse relation between relative profitability and return sensitivity holds regardless of whether the profitability differences among firms arise from cost efficiency or market share.

Further considering the possibility of a capacity limit (an issue ignored in the original IO models), we also predict that the role played by relative profitability in driving cross-firm differences in return sensitivity is asymmetrical between positive and negative industry news. More specifically, the return sensitivity differences induced by relative profitability within a given industry are more pronounced when there is positive rather than negative industry news, especially for industries with a high (versus low) level of capital intensity.

To test these predictions, we rank firms within an industry by profitability (measured by return on equity, ROE) and empirically examine whether a firm's profitability ranking systematically influences the co-movement of its returns with those of the industry (which serve as a proxy for industry-level news). As predicted, we find that the returns of firms that were less profitable in the prior period are significantly more sensitive to industry returns. Furthermore, this inverse relationship between (lagged) relative profitability and return sensitivity is more pronounced in the subsample of positive industry returns than in that of negative industry returns, with the asymmetry between the two types of industry news more prominent for industries with high (versus low) capital intensity. Our results hold after controlling for stock price movements that are attributable to common risk factors and firm-specific accounting information.

Profitability as a summary measure of competiveness reflects strengths or weaknesses in business operations. To examine whether the effects of relative profitability extend to more fundamental factors, we rank firms by cost efficiency and market share and then conduct a similar analysis to that for profitability ranking. We confirm that relative cost efficiency and relative market share each have an effect on return sensitivity similar to that of relative profitability.

Our prediction of an inverse relation between relative profitability and return sensitivity suggests that as a firm's profitability ranking fluctuates within its industry over time, its industry beta (i.e., return sensitivity to industry news) should move in the opposite direction. We thus examine the co-movement of profitability ranking and the industry beta over time and find evidence consistent with our conjecture.

According to our model, earnings behavior is distinctly different from stock return behavior in an industry context. Specifically, the earnings of the more profitable firms in a given industry are more sensitive to industry news than are those of their less profitable counterparts. We also investigate this prediction empirically and produce results consistent with our prediction for mainstream observations (i.e., after excluding extreme outliers in earnings that are likely to be severely distorted and thus unrepresentative of long-term performance).

Finally, we conduct a range of sensitivity tests and demonstrate that our results remain robust to alternative measures of profitability and industry news. They also hold separately in different sub-periods and in subsamples partitioned by industry concentration and firm size. We also produce similar results for single- and multiple-segment firms.

This study contributes to research exploring the fundamental factors driving stock returns. Extant studies have used either aggregate (primarily market-wide) or firm-specific information to explain returns, and there has been little work examining how the two types of information should be integrated. We show that stock returns depend not only on aggregate-level and firm-specific information separately, but also on the interaction between them. Specifically, relative firm profitability within an industry determines the degree to which returns respond to industry news. With a few exceptions such as King (1966) and Magee (1974), the literature has paid limited attention to the role of industry-level information in determining returns.

The study also contributes to earnings response coefficient (ERC) research, which aims to explicate stock price reactions to unexpected earnings. Unexpected earnings can arise from market-wide, industry-wide, or firm-level events. Our study focuses on how prices respond to the component of unexpected earnings that is driven by industry-wide events. It shows that the response (in terms of returns) to such earnings news is greater for the less profitable firms in an industry, thus shedding light on cross-firm variations in ERC arising from particular sources of earnings change. In our empirical analysis, we demonstrate this distinctive effect of industry-level news after controlling for both market-wide and firm-specific information.

By explaining how relative profitability influences return sensitivity, this study also sheds light on inter-firm dependencies in returns. Research on intra-industry information transfer has shown that a firm's earnings news is informative about the performance of other firms in the same industry (e.g., Foster 1981; Han and Wild 1990; Ramash 2002; Thomas and Zhang 2008). In these studies, inter-firm dependencies arise for *informational* reasons under conditions of investor uncertainty. As firms operating in the same industry are subject to common exogenous shocks, the release of one firm's performance will give an indication of the performance of other firms, which results in informational spillover. In contrast, our study considers inter-firm dependencies that arise from firms' strategic decision-making: Each firm makes operational decisions that are conditional on the decisions of other firms in the same industry, thus resulting in an *operational* inter-dependency that exists regardless of whether investors are informed about firm operations.

Finally, this study is relevant to investors. First, as already mentioned, it shows how a firm's profitability ranking in the industry indicates its systematic risk relative to that of its industry peers. Second, our results suggest that, in a market with differentially informed investors, those investors who believe that they have special insight into the future development of a particular industry (and whose view is not yet reflected in market prices) should pick stocks conditional on their outlook for the industry. Investment allocation should favor the less profitable firms in an industry when the outlook is good and switch to the more profitable firms when it is poor. This investment strategy is distinct from the contrarian strategy (e.g., De Bondt and Thaler 1989; Chan 1988; Lakonishok et al. 1994; Dechow and Sloan 1997) that relies on a stock's past performance (rather than industry outlook) to form portfolios.

In Section II, we develop theoretical predictions from a Cournot competition model. Section III describes our empirical methods and data. Sections IV to VIII provide evidence on the model's predictions, and Section IX discusses the robustness of the results. Section X concludes. Appendix A extends the theoretical analysis to other scenarios (including Bertrand competition).

II. THEORETICAL ANALYSIS AND HYPOTHESES

In this section, we develop a model that demonstrates the importance of a firm's profitability relative to its industry peers in determining the sensitivity of its returns to industry-wide shocks. We first establish the role played by relative profitability in a one-period Cournot (quantity) competition model. In Appendix A, we show that our theoretical results can be extended to other scenarios, including alternative causes of profitability differences, and Bertrand (price) competition models.

A Model with Cournot Competition

Consider a duopoly in which two firms producing an identical product engage in Cournot competition.[1] Let C_j be the variable cost of production for firm j, $j = 1,2$. The two firms have the same fixed costs, which are set to zero for simplicity. Without loss of generality, let $C_1 < C_2$, such that firm 1 is more cost-efficient than firm 2.[2] The demand function for the product is, $P = a - bX$, where P is the product price, X is the aggregate demand, and a and b are constants.

Let x_j be the quantity of output produced by firm j ($j = 1,2$). The problem facing each firm is to determine the quantity that maximizes its profit, conditional on that produced by the other firm, that is:

$$\underset{x_j}{Max}[a - b(x_1 + x_2) - C_j]x_j,\ j = 1,2. \tag{1}$$

The first-order conditions are:

$$a - bx_j - b(x_1 + x_2) - C_j = 0,\ j = 1,2. \tag{2}$$

The optimal quantities determined from (2) are:

$$x_j = \frac{1}{3b}[a - 3C_j + (C_1 + C_2)],\ j = 1,2. \tag{3}$$

---

[1] Our qualitative predictions can be generalized to an industry with three or more firms.

[2] In Appendix A, we also consider a scenario in which cost efficiency differences are attributed to fixed costs. It should be noted, however, that fixed costs are less of an issue when a long-term perspective is adopted. In the long run, firms may adjust their capacities, which means that costs that are fixed in the short run become variable.

Substituting (3) into (1) and simplifying, we obtain the equilibrium earnings, E_j, as:

$$E_j = bx_j^2 = \frac{1}{9b}[a - 3C_j + (C_1 + C_2)]^2,\ j = 1,2. \tag{4}$$

Let V_j be the market value of firm j, which for simplicity is assumed to equal the capitalization of its earnings, and r_j be the discount rate appropriate for firm j. Then:

$$V_j = \frac{1}{9br_j}[a - 3C_j + (C_1 + C_2)]^2,\ j = 1,2. \tag{5}$$

To examine the impact of industry-wide news, we now assume an unexpected shock, Δa, that shifts the demand curve to $P = a + \Delta a - bX$. After adjusting the firms' production quantities, their values become:

$$V_j' = \frac{1}{9br_j}[a + \Delta a - 3C_j + (C_1 + C_2)]^2,\ j = 1,2.$$

The stock returns triggered by the demand shock are:[3]

$$R_j \equiv \frac{V_j' - V_j}{Vj} = \frac{2(\Delta a)}{a - 3C_j + (C_1 + C_2)},\ j = 1,2. \tag{6}$$

Several observations can be made about the equilibrium quantities, earnings, and stock returns. First, it is obvious from (3) that $\partial x_j/\partial C_j = -2/(3b) < 0$, indicating that the more profitable firm in the industry (the one with a lower cost) produces a larger quantity of output. Second, from (4), which shows that earnings relate to output quantity squared, we get $\partial E_j/\partial a = 2x_j/3 > 0$ and $\partial(\partial E_j/\partial a)/\partial C_j = -4/(9b) < 0, j = 1,2$. Thus, while a common shock causes the earnings of all firms to change in the same direction, the magnitude of the earnings impact is greater for the firm with higher profitability (or, equivalently, a lower cost). Third, given, $C_1 < C_2$ it follows from (6) that $R_2 > R_1 > 0$ if $\Delta a > 0$, and $R_2 < R_1 < 0$ if $\Delta a < 0$. That is, industry shocks cause firm returns to move in the same direction and the magnitude of the return impact is greater for the firm with lower profitability.[4] Note that (6) can be rewritten as $R_j = 2\Delta a/(3bx_j)$, so that the return magnitude is inversely related to output quantity (which, in turn, is positively related to profitability).

Our finding that more profitable firms capture more earnings gains as industry conditions improve and, conversely, relinquish more earnings as conditions deteriorate, should be intuitive. Less intuitive, perhaps, is the opposite finding we reach for investors, who capture more gains by holding the stocks of less, versus more, profitable firms when there are favorable shocks. To reconcile the different consequences for firms and investors, it should be noted that the impact on a firm's (operational) performance is measured in terms of the earnings change, whereas that on investors is measured in terms of the value change (which arises from the earnings change) *relative to* the initial firm value. Investors pay less for the stock of a less (versus more) profitable firm because of that firm's lower expected earnings. Although a given shock produces a smaller impact on the earnings of a less profitable firm in absolute terms, its proportional impact (i.e., on returns) is actually greater because of the firm's lower initial value.

---

[3] The second-order effect of the shock (which is typically of a much smaller magnitude than the first-order effect) is ignored in the return expressions.

[4] To keep the model simple, we ignore (firm-specific) idiosyncratic news. While a common industry shock impacts the stock prices of different firms in the same direction, idiosyncratic news (being independent across firms) causes firm returns to move apart. Thus, in an extended model that also incorporates idiosyncratic news, for a given industry shock, there will generally be both positive and negative returns in the same industry. Because, by definition, idiosyncratic news is orthogonal to common shocks, its presence does not alter the effect of common shocks, so our predictions concerning return sensitivity to industry-wide news will continue to hold.

The foregoing analysis leads us to our main hypothesis for empirical testing, as follows:

H1: The stock returns of the less profitable firms in a given industry are more sensitive to industry-level news than are the returns of the more profitable firms.

Extensions to Other Scenarios

Our analysis thus far has relied on certain assumptions about firm costs, the form of competition, and the source of profitability differences. In Appendix A, we show that the prediction in H1 holds more generally by extending the analysis to a variety of other scenarios, including (1) differences in fixed costs, (2) shocks occurring on the slope (as opposed to the intercept) of the demand function, and (3) firms engaging in Bertrand competition with differentiated products. In addition, our prediction is also valid in situations where profitability differences are caused by market share and in multiperiod settings (these extensions are available upon request).

Trade-Offs between Variable and Fixed Costs

A limitation of the above model and its extensions in Appendix A is that firms within an industry are assumed to differ in either variable or fixed costs, but not both. More generally, there can be differences in both variable and fixed costs, so relative firm profitability will reflect not only differential efficiency in utilizing resources associated with each type of costs, but also differential effectiveness in trading off between variable and fixed costs.

When one firm is more efficient in both variable and fixed costs than another, our basic prediction (H1) on the role of relative profitability in influencing return sensitivity to industry news will unambiguously hold, as the effects of the two types of costs will lead to the same qualitative prediction. However, in situations in which a firm is relatively efficient in one type of costs but relatively inefficient in the other, the effects of variable and fixed costs will work in opposite directions. In this case, it is not immediately clear whether firms with lower overall profitability will still have returns more sensitively affected by industry news. We leave it to future theoretical research to identify the exact conditions under which H1 is still valid in this more general setting.[5] Nonetheless, to the extent that in the empirical setting, differences in profitability among firms arise mostly from differential efficiency in utilizing resources, rather than from differential effectiveness in trading off between variable and fixed costs, our basic prediction should continue to hold.

Implications of Capacity Limits

The preceding analysis assumes that firms can freely adjust their output to the desired level without having to alter their cost functions. In reality, firms have finite capacities that are fixed in the short term or may be changed only at additional cost. Thus, when faced with favorable shocks, firms may not be able to raise their output levels by as much as the IO models predict.

In the above model, the more profitable firm chooses a higher output level. To the extent that the firms in the model have similar capacities (given that they have the same fixed costs), the more profitable one is likely to reach the capacity limit sooner. Hence, although a favorable shock triggers positive returns for both firms, as previously noted, these positive returns can potentially be dampened because the capacity limits prevent firms from raising their output to the new desired level. Furthermore, the more profitable firm (which already has a higher output level) will be affected more severely by the capacity limit. The consequence of this ceiling effect is greater variation in return sensitivity to industry-wide news across firms with different levels of profitability.[6] We thus predict that relative profitability plays an asymmetrical role in inducing cross-firm differences in return sensitivity in relation to positive and negative industry news, with the effect of relative profitability being more pronounced in regard to the former.

---

[5] In general, firms may differ in the ability to trade-off between variable and fixed costs, and they will choose technologies (cost structures) that reflect their own strengths/weaknesses. The choice of technology in turn affects a firm's ability to respond to positive and negative demand shocks. Again, we leave it to future research to examine how firms strategically choose their cost structures and what implications this has for the link between accounting information and stock returns in an industry context.

The ceiling effect discussed above exists because firms are limited by their existing capacities and are thus unable to quickly adjust output quantities as market conditions improve. To increase capacity, firms must go through a lengthy process of planning investment projects, raising the necessary capital, and installing the assets. Industries differ in the extent of capital investment required for operations, and the problem of investment-caused delays will be more severe in industries requiring high (versus low) levels of capital investment. In other words, capacity constraints will be harder (that is, more difficult to expand in the short term) for firms in more capital-intensive industries, which leads to greater asymmetry in the effect of relative profitability in relation to positive and negative industry news situations. We thus posit Hypotheses 2 and 3, as follows:

H2: The cross-firm variation in relative profitability-induced return sensitivity to industry-wide news is more pronounced when there is positive rather than negative industry news.

H3: The asymmetry in the effects of relative profitability on return sensitivity to positive and negative industry news is more prominent in industries with higher levels of capital intensity.

III. EMPIRICAL METHODS AND SAMPLE DESCRIPTION

In this section, we describe our methods of empirical analysis and the sample.

Principal Empirical Method: Regression Analysis

We use an industry's aggregate returns to proxy for industry news that occurs over a period of time.[7] We employ the following regression model to test whether relative firm profitability affects a firm's return sensitivity to such news:

$$R^i_{k,t} = \alpha + \beta_1 IR^i_t + \beta_2 rp^i_{k,t-1} + \beta_3 IR^i_{k,t} * rp^i_{k,t-1} + \sum \gamma_j Control_j + \varepsilon^i_{k,t} \qquad (7)$$

where $R^i_{k,t}$ represents the return in period t of firm k in industry i; $IR^i_t \equiv \sum_{s \neq k} R^i_{s,t}/(N_i - 1)$ represents the return in period t of industry i, constructed as the equally weighted return of all (N_i)

---

[6] There is also a potential "floor" effect in situations of negative industry shocks, which limits the extent of downward output adjustments. This effect arises because of the various costs involved in reducing a firm's activity level (for example, the costs of layoffs). It is not immediately obvious whether the more profitable firms are more susceptible to such a floor effect than their less profitable counterparts, and so we treat the effect as an empirical issue.

[7] Industry average returns have often been used in prior research to represent industry-wide news and changes in industry conditions (e.g., Roll 1988; Durnev et al. 2003). A limitation of this proxy is that it can be distorted by a few firms' extreme returns (outliers) or there may be instances where there is no clear shift in the overall condition of an industry. In Section IX ("Robustness Checks"), we discuss the additional analyses performed to address these issues.

firms in industry *i*, excluding firm *k* (to avoid spurious relations);[8] and $rp^i_{k,t-1}$ is firm *k*'s relative profitability in industry *i* in the prior period (period *t*−1), defined as the ranking of firm *k*'s ROE in industry *i* in period *t*−1 (in ascending order starting from 1) divided by the number of firms in industry *i* in period *t*−1. Thus, $rp^i_{k,t-1}$ takes a value between 0 and 1 and is greater for the more profitable firms in the industry. We also include various control variables in (7), as explained in the following.

According to our prediction, the returns of less profitable firms are more affected by industry-wide news than those of more profitable firms, which posits a negative coefficient on the interaction term in (7): $\beta_3 < 0$. A firm's total return sensitivity to industry returns is given by $\beta_1 + \beta_3 * rp_{t-1}$, which may be viewed as the firm's industry beta.

Finance research has found that cross-sectional returns can be explained by firm size, systematic market risk, and the book-to-market ratio of equity (e.g., Fama and French 1992, 1993, 1996). Accounting research, in contrast, reports that these returns are driven by accounting variables that convey a firm's current-period performance and changes in its expected future cash flows. For example, in Chen and Zhang's (2007) model, returns are a function of earnings yield, profitability changes, equity capital investment, and changes in growth opportunities and the discount rate. Guided by prior research, we include two groups of control variables in Regression (7): (i) lagged systematic risk (*lagMbeta*), lagged size (*lagSize*), lagged financial leverage (*lagDE*), and the lagged book-to-market ratio of equity (*lagBM*), as proxies for common risk factors; and (ii) earnings yield (*Earn*), profitability change (Δp), equity capital investment (*CapInv*), and the discount rate change (Δr), as firm-specific informational variables.[9] Because the effects of accounting variables have been shown to vary with profitability (e.g., Chen and Zhang 2007), we also include the interactions between these variables and relative profitability. Finally, we include market returns, *MR*, as another control.

The results reported below are based on the ordinary least squares (OLS) procedure, with standard errors corrected for heteroscedasticity and cross and serial correlations. We employ the Fama and MacBeth (1973) procedure as an alternative in our robustness checks. In addition to the regression method, we also report a portfolio approach in Section VI.

Data and Variable Definitions

We retrieve accounting data from the Compustat Quarterly File. For comparability of firm profitability and industry news, we restrict our sample to firms with a reporting quarter ending in December. We treat this quarter as the fourth quarter and combine it with the three preceding quarters to form a fiscal year. Based on this definition of a fiscal year, we construct annual data for the accounting variables,[10] and calculate annual returns using stock prices from the CRSP Monthly Return Files.

We measure the accounting variables as follows. Profitability is measured by ROE, which is defined as earnings before extraordinary items divided by the book value of equity at the beginning of the year; the earnings yield ($Earn_t$) is per-share earnings before extraordinary items divided by the stock price at the beginning of the current year (year *t*); profitability change (Δp_t) equals the current year's ROE minus the prior year's ROE; and equity capital investment ($CapInv_t$) equals the

---

[8] We also employ value-weighted industry returns and find the results qualitatively unchanged.

[9] We exclude one of the variables in Chen and Zhang's (2007) model, growth opportunity changes, for the following reasons: (1) this variable is often proxied by changes in analyst forecasts of long-term earnings growth, and the additional data required substantially reduce the sample size; and (2) the inclusion of this variable has no significant incremental effects in our regressions.

[10] This approach properly matches the time period for comparing the profitability of different firms in an industry. Alternatively, if we follow actual reporting years, then the maximum mismatch in time among firms can be as much as 11 months, which impairs cross-firm comparability.

TABLE 1

Definitions of Main Variables

Variable	Definition
$rp^i_{k,t-1}$	Relative profitability of firm *k* in industry *i* in year *t*−1, calculated as the ranking of firm *k*'s ROE in the industry (in ascending order starting from 1) divided by the number of firms in industry *i* in period *t*−1, where ROE equals earnings before extraordinary items divided by the book value of equity at the beginning of the period;
$R_{k,t}$	Stock return of firm *k* in year *t*;
MR_t	Return on the market portfolio, calculated as the equally weighted average return of all listed firms in year *t*;
IR^i_t	Return of industry *i* in period *t*, calculated as the equally weighted average return of all firms in the industry (excluding the firm whose return is being explained in the regression);
$Earn_t$	Earnings yield of a firm in year *t*, equal to per-share earnings before extraordinary items divided by the share price at the beginning of the period;
Δp_t	Firm profitability (ROE) change in year *t* relative to year *t*−1;
$CapInv_t$	Firm equity capital investment in year *t*, calculated as the change in equity book value over the year normalized by the beginning equity book value;
Δr_t	Discount rate change in year *t* relative to year *t*−1, with the discount rate (*r*) proxied by the yield on the ten-year Treasury Bond;
lagMbeta	A firm's lagged market risk, estimated from the market model based on the returns of the 36 preceding monthly returns;
lagSize	Logarithm of a firm's market value of equity at the end of year *t*−1;
lagBM	Ratio of a firm's book value to market ratio of equity at the end of year *t*−1; and
lagDE	A firm's long-term debt-to-equity ratio at the end of year *t*−1.

change in equity book value over the year divided by the beginning equity book value. The discount rate change (Δr_t) is proxied by the change in the yield on the ten-year Treasury Bond during the current year.

The lagged market beta, *lagMbeta*, is estimated by the capital asset-pricing model (CAPM) on the basis of the preceding 36 monthly returns; lagged size, *lagSize*, is measured as the logarithm of the firm's equity market value at the end of the prior year; the lagged book-to-market ratio of equity, *lagBM*, is measured at the end of the prior year; and the lagged ratio of long-term debt to equity market value, *lagDE*, is also measured at the end of the prior year. Table 1 provides a detailed definition of all variables.

We use four-digit SICs to identify industries. The following procedure was employed for sample selection. We excluded observations in an industry-year with fewer than ten firms and observations with missing data for the variables. The top and bottom 1 percent of extreme values for each of the continuous variables are winsorized. We exclude utility and financial firms as in prior research (e.g., Fama and French 1992; Burgstahler and Dichev 1997; Ryan and Zarowin 2003).[11] Our final sample consists of 43,768 observations of firms in 138 four-digit SIC industries during the 1973–2004 period.

Table 2, Panel A reports descriptive statistics. Annual firm-level returns (*R*) have a mean (median) of 19.3 percent (5.9 percent), industry returns a mean (median) of 20.5 percent (12.2 percent), and

---

[11] In our separate analysis for these industries the results were weaker or inconsistent with the predictions, suggesting that our theoretical predictions, which arise from a purely market-based setting, are less applicable in industries that are subject to more extensive rules and regulations (which limit firms' abilities to freely respond to external economic shocks).

TABLE 2

Descriptive Statistics of Variables and Correlations

Panel A: Descriptive Statistics of Variables in Overall Sample

Variable	Mean	Median	Std. Dev.	Min.	Q1	Q3	Max.
R	0.193	0.059	0.817	−0.954	−0.241	0.400	10.005
MR	0.131	0.176	0.176	−0.278	−0.007	0.284	0.373
IR	0.205	0.122	0.442	−0.699	−0.055	0.349	3.031
rp_{t-1}	0.518	0.524	0.297	0.000	0.267	0.777	1.000
$lagMbeta$	1.149	1.049	0.917	−2.224	0.563	1.606	5.398
$lagSize$	4.823	4.656	2.174	−0.558	3.228	6.310	11.359
$lagBM$	0.796	0.588	0.748	0.023	0.325	1.002	7.917
$lagDE$	0.599	0.243	1.089	0.000	0.012	0.706	10.335
$Earn$	−0.013	0.042	0.247	−2.414	−0.045	0.090	0.895
Δp	−0.017	−0.003	0.407	−4.057	−0.095	0.062	2.816
$CapInv$	0.164	0.072	0.655	−0.911	−0.059	0.208	8.415
Δr	−0.001	−0.001	0.012	−0.031	−0.010	0.006	0.025

Panel B: Correlations

Variable	R	MR	IR	rp_{t-1}	$lagMbeta$	$lagSize$	$lagBM$	$lagDE$	$Earn$	Δp	$CapInv$
MR	0.240***										
IR	0.441***	0.466***									
rp_{t-1}	−0.010**	−0.007	−0.003								
$lagMbeta$	0.057***	−0.026***	0.042***	−0.053***							
$lagSize$	−0.078***	−0.061***	−0.009*	0.310***	0.073***						
$lagBM$	0.107***	0.018***	0.063***	−0.248***	−0.104***	−0.394***					
$lagDE$	0.005	0.005	−0.025***	−0.086***	−0.073***	0.006	0.047***				
$Earn$	0.106***	0.007	−0.013***	0.309***	−0.104***	0.223***	−0.223***	−0.019***			
Δp	0.118***	−0.018***	0.021***	−0.211***	0.002	0.005	0.002	−0.008	0.280***		
$CapInv$	0.258***	0.039***	0.094***	0.039***	0.054***	0.033***	−0.205***	0.115***	0.174***	0.039***	
Δr	0.054***	−0.119***	0.106***	−0.006	−0.010**	0.005	0.034***	0.008	0.067***	0.025***	0.020***

(continued on next page)

TABLE 2 (continued)

Panel C: Distribution of *N* (Number of Firms in an Industry-Year) and *Consistent_R* (Percentage of Firms with Returns in the Same Direction as the Industry Return in an Industry-Year)

	Mean	Median	Std. Dev.	Min.	Q1	Q3	Max.
N	23.658	15.000	22.092	10.000	12.000	23.000	205.000
Consistent_R	0.663	0.680	0.189	0.021	0.545	0.800	1.000

---

*, **, *** Indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

This table reports the distributional statistics of the following variables (defined in Table 1) for the overall sample, in Panel A, and the correlations among the variables, in Panel B: annual firm-level stock return (*R*), market return (*MR*), industry return (*IR*), prior year's relative profitability (rp_{t-1}), lagged market beta (*lagMbeta*), lagged firm size (*lagSize*), lagged book-to-market ratio of common equity (*lagBM*), lagged debt-to-equity ratio (*lagDE*), earnings yield (*Earn*), profitability change (Δp), capital investment (*CapInv*), and discount rate change (Δr). Panel C reports the distributional statistics of two industry-level variables: the number of firms in an industry-year (*N*) and the percentage of firms having returns in the same direction as the industry return in an industry-year (*Consistent_R*). The sample consists of 43,768 firm-year observations for the 1973–2004 period (32 years) involving 138 four-digit SIC industries.

CALCUTTA UNIVERSITY

market returns a mean (median) of 13.1 percent (17.6 percent). The standard deviation is 81.7 percent for firm-level returns, 44.2 percent for industry returns, and 17.6 percent for market returns, which indicates that returns are much more volatile at the firm level than at the aggregate levels.

The means (medians) of the control variables are 1.149 (1.049) for *lagMbeta*, 4.823 (4.656) for *lagSize*, 0.796 (0.588) for *lagBM*, and 0.599 (0.243) for *lagDE*. The mean (median) is −0.013 (0.042) for the earnings yield (*Earn*); −0.017 (−0.003) for the profitability change (Δp); 0.164 (0.072) for capital investment (*CapInv*); and −0.001 (−0.001) for the discount rate change (Δr).

Table 2, Panel B presents the pairwise Pearson correlations between the variables, which offers a preliminary view of how firm returns are related to both industry returns and the firm-specific variables. First, we observe that a firm's returns are highly correlated with both industry and market returns, although this correlation is substantially higher with industry returns (0.441) than with market returns (0.240).

Second, firm returns are also highly correlated with the accounting variables, including the earnings yield (the correlation equals 0.106), profitability change (0.118), and equity capital investment (0.258), consistent with the findings of Chen and Zhang (2007). In addition, returns are correlated with firm characteristics that are commonly viewed as risk proxies, including the book-to-market ratio (the correlation equals 0.107), size (−0.078), and market beta (0.057), consistent with Fama and French (1992, 1993, 1996).

Correlation analysis also reveals that, on average, highly profitable firms are large in size and have a low book-to-market ratio (consistent with the predictions of Zhang [2000]) and a low degree of leverage. A high level of firm profitability predicts a high earnings yield in the subsequent period, but a decline in profitability, consistent with prior findings that profitability is mean-reverting (e.g., Dechow et al. 1999; Biddle et al. 2001). A high level of profitability is also followed by a high level of equity capital investment, consistent with Zhang (2000) and Biddle et al. (2001).

Table 2, Panel C shows the distributions of two industry-level variables: the number of firms (*N*) and the percentage of firm returns having the same sign as that of the industry return (*Consistent_R*) in an industry-year. *N* has a mean (median) value of 23.66 (15), and *Consistent_R* has a mean (median) value of 66.3 percent (68.0 percent), indicating that, on average, around two-thirds of the firms in an industry have returns moving in the direction of the industry return.

IV. REGRESSION RESULTS FOR H1 AND H2

Relative Profitability and Return Sensitivity to Industry News

In examining the relation between relative profitability and return sensitivity, we initially exclude the control variables in (7) to allow assessment of the "complete" effects of industry news and its interaction with (lagged) relative profitability. As Table 2 shows, the control variables are significantly correlated with industry returns (*IR*), raising the possibility that the variables that are contemporaneous to *IR* may in fact be driven by *IR*. If this is the case, then including these variables would dilute the effect of *IR*.

Table 3, Panel A presents the results of the regression of firm-level returns (R_t) on the prior year's relative profitability (rp_{t-1}), industry returns (IR_t), and the interaction between the two ($IR_t * rp_{t-1}$). In the overall sample, the coefficients on all three variables are significant at the 0.01 level. The coefficient on *IR* is 1.192 (t = 29.46), and that on $IR * rp_{t-1}$ is −0.708 (t = −12.49). Thus, on average, the returns of firms ranked low in terms of the prior year's profitability (rp_{t-1} close to 0) fluctuate more than the corresponding industry returns. In contrast, for firms ranked at the top in terms of profitability ($rp_{t-1} = 1$) the combined coefficient of *IR* is 0.484, which indicates that, on average, their returns are (1) less volatile than industry returns and (2) much less sensitive to industry news relative to their low-profitability counterparts. These results show that, although

TABLE 3

Relative Firm Profitability and Return Sensitivity to Industry News: Overall Sample

Panel A: Results from the Overall Sample without Controls

Variable	**Combined Sample (A)**	**Positive Industry News (B)**	**Negative Industry News (C)**
Constant	−0.040***	−0.072***	−0.033**
	(−5.63)	(−4.18)	(−2.28)
IR	1.192***	1.241***	1.067***
	(29.46)	(21.86)	(19.49)
rp_{t-1}	0.123***	0.232***	0.112***
	(11.32)	(9.34)	(4.86)
$IR * rp_{t-1}$	−0.708***	−0.871***	−0.251***
	(−12.49)	(−11.05)	(−2.80)
Diff. in $IR * rp_{t-1}$: *(B) − (C)*		−0.620***	
(χ^2-statistic)[a]		(28.03)	
Obs.	43,768	29,865	13,903
Adj. R^2	0.208	0.147	0.087

Panel B: Results from the Overall Sample with Controls

Variable	**Combined News (A)**	**Positive Industry News (B)**	**Negative Industry News (C)**
Constant	−0.200***	−0.278***	−0.108***
	(−11.96)	(−11.06)	(−5.30)
IR	1.065***	1.119***	0.783***
	(27.25)	(21.38)	(14.17)
rp_{t-1}	0.199***	0.324***	0.111***
	(13.79)	(11.92)	(4.92)
$IR * rp_{t-1}$	−0.660***	−0.799***	−0.226***
	(−12.46)	(−11.08)	(−2.66)
MR	0.214***	0.154***	0.212***
	(10.13)	(5.26)	(8.52)
lagMbeta	0.049***	0.089***	−0.041***
	(10.04)	(12.72)	(−9.00)
lagSize	−0.021***	−0.027***	−0.002
	(−12.19)	(−11.80)	(−1.08)
lagBM	0.149***	0.159***	0.121***
	(18.24)	(15.66)	(14.41)
lagDE	−0.008**	−0.003	−0.021***
	(−2.02)	(−0.46)	(−5.58)
Earn	0.201***	0.214***	0.237***
	(5.16)	(4.25)	(6.46)
$Earn * rp_{t-1}$	0.326***	0.300***	0.281***
	(3.87)	(2.77)	(3.31)
Δp	0.163***	0.173***	0.134***
	(9.41)	(7.24)	(9.45)
$\Delta p * rp_{t-1}$	−0.001	0.002	0.006
	(−0.45)	(0.38)	(0.45)

(continued on next page)

TABLE 3 (continued)

Variable	Combined News (A)	Positive Industry News (B)	Negative Industry News (C)
CapInv	0.279***	0.319***	0.143***
	(18.63)	(15.50)	(11.22)
$CapInv * rp_{t-1}$	−0.005	0.011	0.001
	(−0.40)	(0.47)	(0.07)
Δr	0.142	0.188	−0.429
	(0.51)	(0.51)	(−1.36)
Diff. in $IR * rp_{t-1}$: *(B) − (C)*			−0.573***
(χ^2-statistic)[a]			(27.03)
Obs.	43,768	29,865	13,903
Adj. R^2	0.294	0.248	0.214

, * Indicate significance at the 0.05 and 0.01 levels, respectively.

[a] We employ the seemingly unrelated estimation (SUE) procedure to test for cross-model differences in the coefficient on $IR * rp_{t-1}$.

This table reports the results of the following regression:

$$R^i_{k,t} = \alpha + \beta_1 IR^i_t + \beta_2 rp^i_{k,t-1} + \beta_3 IR^i_{k,t} * rp^i_{k,t-1} + \sum \gamma_j Control_j + \varepsilon^i_{k,t} \quad (7)$$

where $R^i_{k,t}$ represents the return in period t of firm k belonging to industry i; $IR^i_t = \sum_{s \neq k} R^i_{s,t}/(N_i - 1)$ is the return in period t of industry i, calculated as the equally weighted average return of all firms in industry i, excluding firm k, and $rp^i_{k,t-1}$ is firm k's relative profitability in industry i in the prior period (period t−1). The control variables are market returns (*MR*), lagged market beta (*lagMbeta*), lagged firm size (*lagSize*), lagged book-to-market ratio of equity (*lagBM*), lagged debt-to-equity ratio (*lagDE*), earnings yield (*Earn*), profitability change (Δp), capital investment (*CapInv*), and discount rate change (Δr). The variable definitions are provided in Table 1. Positive industry news is proxied by positive industry returns, and negative industry news by negative industry returns. The sample consists of 43,768 firm-year observations for the 1973–2004 period (32 years), involving 138 four-digit SIC industries. Panels A and B report the results without and with the control variables, respectively. We adopt the OLS procedure, and correct standard errors for heteroscedasticity and cross and serial correlations of the residuals. The numbers in parentheses are t-statistics, unless indicated otherwise.

industry news is an important driver of the returns of individual firms, return sensitivity to such news varies substantially across the firms within an industry, with less profitable firms exhibiting significantly greater sensitivity, which is consistent with H1.

Next, we partition the overall sample into positive and negative industry news, as proxied by positive and negative industry returns, respectively. We find a significantly negative coefficient on $IR_t * rp_{t-1}$ in both subsamples, indicating that relative profitability has a negative effect on return sensitivity to industry news, irrespective of the direction of that news. The coefficient on $IR_t * rp_{t-1}$ is greater in magnitude for the positive industry return subsample, equal to −0.871 ($t = -11.05$; $p < 0.01$), than for the negative subsample, equal to −0.251 ($t = -2.80$; $p < 0.01$), with the difference between them significant at the 0.01 level ($\chi^2 = 28.03$). This indicates that relative profitability better explicates return sensitivity in situations of favorable industry news, consistent with H2.

Table 3, Panel B reports the results after controlling for common risk factors (*lagMbeta*, *lagSize*, *lagDE*, and *lagBM*) and firm-specific variables (*Earn*, Δp, *CapInv*, and Δr). The inverse relation between relative profitability and return sensitivity to industry news continues to hold. The coefficient on $IR_t * rp_{t-1}$ remains negative, and is significant at the 0.01 level in both the overall sample and the subsamples of positive and negative industry news. In fact, the inclusion of the control variables has little effect on the coefficients on the main variables, $IR * rp_{t-1}$ and IR, which indicates that the aforementioned effect of relative profitability is not spuriously caused by the omission of risk factors and firm-specific information. The difference in the coefficient on $IR * rp_{t-1}$

between the positive and negative industry news subsamples is significant at the 0.01 level. These results are consistent with H1 and H2.

We also find that market beta and the book-to-market ratio significantly affect returns in directions that are consistent with prior findings, although the effect of size is significant only in the subsample of positive industry news and that of financial leverage is significant only in the subsample of negative industry news. Earnings yield has a positive effect on returns, with the effect increasing with profitability. Changes in profitability and capital investment both have a positive effect, but these effects do not vary significantly with profitability.

Effects of the Fundamentals: Cost Efficiency and Market Share

Profitability as an indicator of competiveness reflects the fundamentals of business operations. We now examine two fundamental factors, namely, cost efficiency and market share, to determine whether each has an effect on return sensitivity similar to that of relative profitability. We rank the firms in an industry simultaneously by market share (equal to a firm's sales divided by total sales in the four-digit SIC industry) and cost efficiency (equal to earnings before extraordinary items divided by sales) into quintiles on each dimension,[12] and then run Regression (8), which replaces relative profitability (rp) in (7) with two separate variables, relative cost efficiency (rc) and relative market share (rm):

$$\begin{aligned} R^i_{k,t} = {} & \alpha + \beta_1 IR^i_t + \beta_2 rc^i_{k,t-1} + \beta_3 IR^i_{k,t} * rc^i_{k,t-1} + \beta_4 rm^i_{k,t-1} + \beta_5 IR^i_{k,t} * rm^i_{k,t-1} \\ & + \sum \gamma_j Control_j + \varepsilon^i_{k,t} \end{aligned} \tag{8}$$

where rc_{t-1} and rm_{t-1} are the quintile rankings of a firm's cost efficiency and market share in an industry in the prior year, normalized to values between 0 (low cost efficiency or market share) and 1 (high). In our sample, the Pearson (Spearman) correlation between rc_{t-1} and rp_{t-1} is 0.791 (0.788), that between rm_{t-1} and rp_{t-1} is 0.306 (0.301), and that between rc_{t-1} and rm_{t-1} is 0.270 (0.263), all significant at the 0.01 level.

Table 4 shows that the coefficients on $IR_t * rc_{t-1}$ and $IR_t * rm_{t-1}$ are both negative, with the magnitude significantly greater in the subsample of positive industry news than in that of negative industry news. This finding demonstrates that the qualitative effect of relative profitability on return sensitivity extends to relative cost efficiency and relative market share, providing further support for our predictions in H1 and H2.

V. CROSS-INDUSTRY DIFFERENCES IN THE EFFECT OF CAPACITY CONSTRAINTS (H3)

As discussed in Section II, capacity constraints that limit upward output adjustments render the effect of relative profitability on return sensitivity stronger when there is positive rather than negative industry news. We now test H3, which predicts that this asymmetry between positive and negative industry news is more prominent in industries that are more capital intensive and, hence, face harder (e.g., more rigid) constraints. We divide industries into high and low capital intensity groups, with capital intensity measured by the ratio of plant, property, and equipment (PPE) to the number of employees.[13]

Table 5 reports the results for the high and low capital intensity groups separately. In the high capital intensity group, the coefficient on $IR * rp_{t-1}$ is -1.135 conditional on positive industry news

---

[12] We obtain similar results using conditional sorting starting with either cost efficiency or market share.

[13] This measure of capital intensity is commonly adopted in the economics literature (e.g., Leontief 1954; Baldwin 1971; Winston 1979; Arai 2003).

TABLE 4

Relative Cost Efficiency and Market Share and Return Sensitivity to Industry News

Variable	Combined News (A)	Positive Industry News (B)	Negative Industry News (C)
$IR * rc_{t-1}$	−0.400***	−0.461***	−0.176***
	(−9.13)	(−7.45)	(−3.32)
$IR * rm_{t-1}$	−0.315***	−0.383***	−0.054
	(−7.14)	(−6.24)	(−1.07)
Diff. in $IR * rc_{t-1}$: *(B) − (C)*		−0.285***	
(χ^2-*statistic*)[a]		(12.68)	
Diff. in $IR * rm_{t-1}$: *(B) − (C)*		−0.329***	
(χ^2-*statistic*)[a]		(18.22)	
Obs.	41,949	28,723	13,226
Adj. R^2	0.297	0.263	0.248

*** Indicates significance at the 0.01 level.

[a] We employ the seemingly unrelated estimation (SUE) procedure to test for cross-model differences in the coefficient on $IR * rp_{t-1}$.

This table reports the results based on the following regression:

$$R^i_{k,t} = \alpha + \beta_1 IR^i_t + \beta_2 rc^i_{k,t-1} + \beta_3 IR^i_{k,t} * rc^i_{k,t-1} + \beta_4 rm^i_{k,t-1} + \beta_5 IR^i_{k,t} * rm^i_{k,t-1} + \sum \gamma_j Control_j + \varepsilon^i_{k,t} \quad (8)$$

where $R^i_{k,t}$ represents the return in period t of firm k belonging to industry i; $IR^i_t = \sum_{s \neq k} R^i_{s,t}/(N_i - 1)$ is the return in period t of industry i, calculated as the equally weighted average return of all firms in industry i, excluding firm k; $rc^i_{k,t-1}$ is firm k's relative cost efficiency in industry i in period t−1, and $rm^i_{k,t-1}$ is firm k's relative market share in industry i in period t−1, both of which are measured by within-industry quintile rankings. We perform independent sorting on cost efficiency (earnings before extraordinary items divided by sales) and market share (firm sales divided by total sales in the four-digit SIC industry) to obtain these rankings, with lower values denoting lower cost efficiency or a smaller market share. The definitions of the other variables are provided in Table 1. Positive industry news is proxied by positive industry returns, and negative industry news by negative industry returns. The sample consists of 41,949 firm-year observations for the 1973–2004 period (32 years). We adopt the OLS procedure, and correct standard errors for heteroscedasticity and cross and serial correlations of the residuals. The numbers in parentheses are t-statistics, unless indicated otherwise. The coefficients on control variables are omitted for brevity.

and −0.233 conditional on negative industry news. The difference between the two coefficients is significant at the 0.01 level. In the low capacity intensity group, the coefficient on $IR * rp_{t-1}$ is −0.798 conditional on positive industry news and −0.351 conditional on negative industry news. The difference between the two coefficients is significant at the 0.05 level. Thus, the results for each group alone are consistent with H1 and H2.

Importantly, the degree of asymmetry in the effect of relative profitability differs between the two subsamples. In the high capital intensity group, in which the asymmetry is expected to be more pronounced, the difference in the coefficient on $IR * rp_{t-1}$ is −0.902 between positive and negative industry news. In contrast, in the low capital intensity group, in which the asymmetry is expected to be less pronounced, the difference in the coefficient on $IR * rp_{t-1}$ between the two types of news is −0.447. The difference in the extent of asymmetry is significant at the 0.01 level between the high and low capital intensity groups. These results are consistent with H3.

VI. SUPPLEMENTARY ANALYSIS FOR H1: THE PORTFOLIO APPROACH

To supplement the above regression analysis, we form annual portfolios by assigning the firms in each industry into deciles based on the prior year's profitability (ROE) and combining firms of

TABLE 5

Effects of Capacity Constraints in the High and Low Capital Intensity Industries

	High Capital Intensity		**Low Capital Intensity**	
Variable	**Positive Industry News (A)**	**Negative Industry News (B)**	**Positive Industry News (C)**	**Negative Industry News (D)**
Constant	−0.192***	−0.060*	−0.160***	−0.024
	(−4.27)	(−1.84)	(−4.19)	(−0.59)
IR	1.398***	0.851***	1.208***	0.959***
	(15.12)	(10.19)	(14.24)	(10.09)
Rp_{t-1}	0.328***	0.102***	0.134***	0.035
	(6.48)	(3.06)	(2.87)	(0.75)
$IR * rp_{t-1}$	−1.135***	−0.233*	−0.798***	−0.351**
	(−8.92)	(−1.81)	(−6.65)	(−2.35)
MR	0.111**	0.142***	0.176***	0.144***
	(2.22)	(3.64)	(3.40)	(3.33)
lagMbeta	0.101***	−0.040***	0.115***	−0.040***
	(7.64)	(−4.91)	(10.67)	(−4.91)
lagSize	−0.032***	−0.004	−0.038***	−0.002
	(−7.99)	(−1.37)	(−9.13)	(−0.59)
lagBM	0.100***	0.104***	0.111***	0.100***
	(6.66)	(6.25)	(6.89)	(5.70)
lagDE	−0.000**	−0.008***	−0.003	−0.003
	(−2.15)	(−2.89)	(−0.63)	(−0.82)
Earn	0.129	0.116***	0.207***	0.235***
	(1.19)	(4.62)	(4.72)	(2.66)
$Earn * rp_{t-1}$	0.354*	0.310**	0.839***	0.581**
	(1.76)	(2.16)	(4.09)	(2.35)
Δp	0.013	−0.006	0.041*	−0.000
	(1.23)	(−0.54)	(1.69)	(−0.00)
$\Delta p * rp_{t-1}$	−0.013	0.133***	0.136*	0.206***
	(−0.66)	(2.91)	(1.94)	(2.96)
CapInv	0.020	0.047***	0.014	0.023
	(0.98)	(4.21)	(0.90)	(1.58)
$CapInv * rp_{t-1}$	0.215**	−0.002	0.430***	0.147***
	(2.50)	(−0.16)	(7.80)	(2.89)
Δr	0.103	0.390	0.258	−1.216**
	(0.17)	(0.78)	(0.40)	(−2.47)
Diff. in $IR * rp_{t-1}$: *(A) − (B); (C) − (D)*	−0.902***		−0.447**	
(χ^2*-statistic*)	(25.93)		(5.27)	
Diff. in $IR * rp_{t-1}$: [(*A*) − (*B*)] − [(*C*) − (*D*)]	−0.455***			
(χ^2-statistic)	(6.60)			
Obs.	11,565	5,814	12,670	6,252
Adj. R^2	0.232	0.170	0.228	0.186

*, **, *** Indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

$$R^i_{k,t} = \alpha + \beta_1 IR^i_t + \beta_2 rp^i_{k,t-1} + \beta_3 IR^i_{k,t} * rp^i_{k,t-1} + \sum \gamma_j\, Control_j + \varepsilon^i_{k,t} \quad (7)$$

(continued on next page)

TABLE 5 (continued)

where $R^i_{k,t}$ represents the return in period t of firm k belonging to industry i; $IR^i_t = \sum_{s \neq k} R^i_{s,t}/(N_i - 1)$ represents the return in period t of industry i, calculated as the equally weighted average return of all firms in industry i, excluding firm k; and $rp^i_{k,t-1}$ is firm k's relative profitability in industry i in the prior period (period $t-1$). The control variables are market returns (*MR*), lagged market beta (*lagMbeta*), lagged firm size (*lagSize*), lagged book-to-market ratio of equity (*lagBM*), lagged debt-to-equity ratio (*lagDE*), earnings yield (*Earn*), profitability change (Δp), capital investment (*CapInv*), and discount rate change (Δr). The variable definitions are provided in Table 1. We divide the industries into groups of high and low capital intensity, measured by the ratio of PPE to the number of employees. Positive industry news is proxied by positive industry returns, and negative industry news by negative industry returns. We adopt the OLS procedure, and correct standard errors for heteroscedasticity and cross and serial correlations of the residuals. The numbers in parentheses are t-statistics, unless indicated otherwise. This table reports the results of Regression (7) for industries with high and low capital intensity. We employ the seemingly unrelated estimation (SUE) procedure to test for cross-model differences in the coefficient on $IR * rp_{t-1}$.

the same decile ranking across industries. We employ the intercept (α) of the following four-factor model to measure portfolio performance:[14]

$$R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 Momentum_t + \varepsilon_t, \tag{9}$$

where $(R_t - R_f)$ is the return on a portfolio in year t minus the risk-free rate in year t, $RMRF_t$ is the return on the equally weighted market portfolio minus the risk-free rate, and SMB_t (small minus big), HML_t (high minus low), and $Momentum_t$ are the returns in year t on zero-investment, factor-mimicking portfolios designed to capture size, book-to-market, and momentum effects, respectively. The return accumulation period begins four months after the end of the prior fiscal year (i.e., the year in which ROE is measured).

Table 6 reports portfolio performance by profitability deciles. In the positive industry return periods, the less profitable portfolios (within industries) generally perform better. The abnormal return, as captured by α in the four-factor model (9), shows a generally declining trend from the low to high relative profitability deciles. The Spearman rank correlation between the profitability decile and α is $-$ 0.936, significant at the 0.01 level, thus confirming that firms with a lower level of profitability relative to their industry peers perform better in periods of positive industry returns. The hedge portfolio of a long position in profitability decile 1 and a short position in decile 10 earns an annual abnormal return of 0.245 (which would be available for an investor who is able to predict the industry-wide trend), significant at the 0.01 level.

The opposite situation occurs in periods of negative industry returns. The abnormal return exhibits a generally rising trend from the low to high relative profitability deciles. The Spearman rank correlation between the relative profitability decile and α is 0.564, significant at the 0.1 level, which indicates that firms with a low level of profitability relative to their industry peers perform worse in periods of negative industry news. The hedge portfolio of a long position in profitability decile 10 and a short position in decile 1 earns an annual abnormal return of 0.086, significant at the 0.05 level. Together, these results are consistent with H1.[15]

VII. TIME SERIES CO-MOVEMENT OF RELATIVE PROFITABILITY AND RETURN SENSITIVITY

The preceding analysis shows that a firm's return sensitivity (industry beta) is inversely related to its relative profitability within the industry. In this section, we adopt a time-series perspective to further examine this relation. The specific question we consider is this: As a firm's profitability

---

[14] The results are similar if we use size-adjusted returns or Jensen's alpha to measure performance.

[15] Fama and French (2008) examine abnormal returns on investment strategies based on firm profitability (as opposed to relative profitability within an industry) that are unconditional (as opposed to conditional) on future industry returns.

TABLE 6

Portfolio Returns by Relative Profitability Deciles

	Positive Industry News		**Negative Industry News**	
***rp*-decile**	**Median ROE**	**α**	**Median ROE**	**α**
lowest	−0.186	0.309***	−0.243	−0.155***
		(4.91)		(−5.24)
2	−0.027	0.212***	−0.061	−0.123***
		(2.78)		(−5.40)
3	0.036	0.244***	0.016	−0.085***
		(3.93)		(−3.63)
4	0.072	0.163***	0.056	−0.063***
		(4.56)		(−2.70)
5	0.097	0.134***	0.085	−0.107***
		(3.46)		(−4.60)
6	0.121	0.150***	0.115	−0.058***
		(4.97)		(−3.08)
7	0.150	0.142***	0.148	−0.033
		(4.66)		(−1.30)
8	0.177	0.134***	0.179	−0.067***
		(3.82)		(−3.89)
9	0.222	0.097***	0.224	−0.072***
		(2.79)		(−3.77)
highest	0.313	0.064	0.337	−0.069***
		(1.56)		(−3.51)
Hedge return		0.245***		0.086**
		(3.06)		(2.44)
Spearman rank correlation		−0.936***		0.564*

*, **, *** Indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Portfolios are formed annually by assigning firms into decile groups on the basis of the prior year's profitability (ROE) rankings within industries and then combining firms of the same decile ranking across industries. Portfolio performance is measured by the alpha (α) of the following four-factor model:

$$(R_{pt} - R_{ft}) = \alpha_p + \beta_{1p}RMRF_t + \beta_{2p}SMB_t + \beta_{3p}HML_t + \beta_{4p}Momentum_t + \varepsilon_{pt} \qquad (9)$$

where R_{pt} denotes the returns on portfolio p in year t, R_{ft} is the risk-free rate measured by the annual yield on the T-bill, and $RMRF$ is equally weighted market returns minus the risk-free rate. SMB_t (small minus big), HML_t (high minus low), and $Momentum_t$ are the annual returns on zero-investment, factor-mimicking portfolios designed to capture size, book-to-market, and momentum effects, respectively. The return accumulation period begins from four months after the end of the prior fiscal year in which ROE is measured. The hedge portfolio conditional on positive (negative) industry news takes a long (short) position in the lowest *rp* decile and a short (long) position in the highest *rp* decile. The Spearman rank correlation is between the *rp*-decile rankings and alpha rankings. Positive industry news is proxied by positive industry returns, and negative industry news by negative industry returns. We adopt the OLS procedure, and correct standard errors for heteroscedasticity and cross and serial correlations of the residuals. The numbers in parentheses are t-statistics.

ranking within the industry fluctuates over time, does its industry beta (*Ibeta*) move in a corresponding fashion, as suggested by H1?

Our analysis proceeds in the following steps. First, for each year, we sort the firms in an industry into profitability deciles and then employ the variable $UP_{k,t}$ ($DOWN_{k,t}$) to indicate the direction of movement in firm k's profitability ranking; it is set to 1 if firm k's profitability decile ranking within the industry has moved up (down) in year t relative to year $t-1$, and 0 otherwise. Second, for each year, we estimate an (average) industry beta for the firms within a decile group

($Ibeta_{k,t}$), which is the slope coefficient on IR_t^i in the regression $R_{k,t}^i = \alpha + \beta IR_t^i + \sum \gamma_j Control_j + \varepsilon_{k,t}^i$, where $R_{k,t}^i$ is firm k's returns in industry i in period t, and IR_t^i is the equally weighted returns of all firms in industry i (excluding firm k) in year t. The control variables are the lagged market beta (*lagMbeta*), lagged firm size (*lagSize*), lagged book-to-market ratio of equity (*lagBM*), lagged debt-to-equity ratio (*lagDE*), earnings yield (*Earn*), profitability change (Δp), and capital investment (*CapInv*). Third, we run the following regression:

$$\Delta Ibeta_{k,t} = \alpha + \beta_1 UP_{k,t} + \beta_2 DOWN_{k,t} + \varepsilon_{k,t} \tag{10}$$

where $\Delta Ibeta_{k,t} \equiv Ibeta_{k,t} - Ibeta_{k,t-1}$ is the change in firm k's industry beta (proxied by the average *Ibeta* of all firms in the decile) in year t relative to year t–1. According to H1, the coefficient on *UP* should be negative, and that on *DOWN* should be positive.

Table 7 reports the results. For the overall sample, we obtain a negative coefficient on *UP* and a positive coefficient on *DOWN*, both significant at the 0.01 level. This indicates that, on average, firms with profitability rankings that move upward from one year to the next experience a decline in their industry beta, and vice versa.

We then partition the observations on the basis of whether the prior year's profitability was above or below the industry median for that year, and then conduct separate analyses. The results are qualitatively the same as those for the overall sample, which is consistent with H1.

VIII. RELATIVE PROFITABILITY AND EARNINGS SENSITIVITY TO INDUSTRY EARNINGS

The above results pertain to equity return behavior in response to industry news. Returns are ultimately driven by fundamental economic activities and, thus, in this section, we explore whether

TABLE 7

Time-Series Co-Movement of Relative Profitability and Industry Beta

Variable	**Overall Sample**	**Subsample with Profitability Above Industry Median in Prior Year**	**Subsample with Profitability Below Industry Median in Prior Year**
Constant	0.016***	0.037***	−0.013***
	(5.57)	(14.70)	(−4.53)
UP	−0.014***	−0.006**	−0.017***
	(−5.72)	(−2.34)	(−7.09)
DOWN	0.021***	0.014***	0.039***
	(9.23)	(4.79)	(16.91)
Obs.	43,768	23,397	20,371
Adj. R^2	0.013	0.005	0.028

, * Indicate significance at the 0.05 and 0.01 levels, respectively.

This table reports the results of the following regression:

$$\Delta Ibeta_{k,t} = \alpha + \beta_1 UP_{k,t} + \beta_2 DOWN_{k,t} + \varepsilon_{k,t} \tag{10}$$

where $UP_{k,t}$ is an indicator variable set to 1 if the decile ranking of firm k's profitability in the industry moves up in year t relative to year t–1, and to 0 otherwise, $DOWN_{k,t}$ is an indicator variable set to 1 if the decile ranking of firm k's profitability in the industry moves down in year t from year t–1, and 0 otherwise, and $\Delta Ibeta_{k,t} = Ibeta_{k,t} - Ibeta_{k,t-1}$, with $Ibeta_{k,t}$ (the industry beta) the coefficient on IR_t^i in the following regression estimated for each relative profitability decile group in each year: $R_{k,t}^i = \alpha + \beta IR_t^i + \sum \gamma_j Control_j + \varepsilon_{k,t}^i$, where $R_{k,t}^i$ represents the return in period t of firm k belonging to industry i, and IR_t^i is the equally weighted returns of all firms in industry i (excluding firm k). The control variables are the lagged market beta (*lagMbeta*), lagged firm size (*lagSize*), lagged book-to-market ratio of equity (*lagBM*), lagged debt-to-equity ratio (*lagDE*), earnings yield (*Earn*), profitability change (Δp), and capital investment (*CapInv*). The numbers in parentheses are t-statistics.

earnings behavior is also consistent with the model's predictions. According to the theoretical model in Section II, earnings behavior in the industry context is distinctly different from return behavior; specifically, firms with high profitability not only capture more (incremental) earnings than firms with low profitability in the event of positive industry shocks, but also relinquish more earnings in the event of negative shocks. We employ the following regression to examine the extent to which this prediction about earnings behavior holds:

$$\Delta E^i_{k,t} = \alpha + \beta_1 \Delta IE^i_t + \beta_2 rp^i_{k,t-1} + \beta_3 \Delta IE^i_{k,t} * rp^i_{k,t-1} + \varepsilon^i_{k,t} \tag{11}$$

where $\Delta E^i_{k,t}$ (earnings change) is measured as the change in the operating income after depreciation of firm k in industry i in year t relative to year $t-1$, scaled by the firm's operating assets at the beginning of period t (with operating assets defined as total assets minus cash and short-term investments); and ΔIE^i_t is the change in operating income after the depreciation of industry i in year t relative to year $t-1$, scaled by that industry's total operating assets at the beginning of year t. Industry operating income is defined as the sum of the operating income of all firms in the industry, and industry operating assets are defined analogously.

In the theoretical context, the earnings variable is a "pure" performance measure that is undistorted by the way in which earnings are calculated. In the empirical context, however, reported earnings can be highly distorted for some firms because of biased accounting rules and, in certain situations, intentional managerial manipulation. To mitigate the extent of accounting distortion and noise, in this section we exclude observations with extreme outliers in reported performance (described in the following). In addition, we also use operating income in Regression (11), which is a more precise measure of a firm's performance in its main businesses than net income. Correspondingly, we employ operating profitability (defined as operating income after depreciation divided by beginning operating assets), instead of ROE, to rank profitability within an industry.[16]

Table 8 presents the results. In Panel A, we limit the sample to observations with operating profitability between -1.0 and $+1.0$ and further exclude the top and bottom 10 percent of the ΔE distribution. In the combined news group, we find that $\Delta IE * rp_{t-1}$ has a positive coefficient of 0.259 ($t = 6.78$), significant at the 0.01 level, which indicates that the earnings of more profitable firms co-move more closely with industry-level earnings. Furthermore, this relation holds in both the positive and negative industry news subsamples. In Panel B, we limit the sample to observations with operating profitability between -0.75 and $+0.75$ and exclude the top and bottom 10 percent of the ΔE distribution. We find a similar effect of rp on a firm's earnings sensitivity to industry news, with the coefficient on $\Delta IE * rp_{t-1}$ equal to 0.262 ($t = 6.87$; $p < 0.01$). These results provide additional evidence consistent with the theoretical model.

IX. ROBUSTNESS CHECKS

To check the robustness of our results, we perform a range of sensitivity tests, the results of which are untabulated but are available upon request.

Effects of Firm Size versus Relative Profitability

To ensure that the effect found for relative profitability is not caused spuriously by firm size, we conduct two tests. First, we divide firms into small- and large-size subsamples and perform

---

[16] The inclusion of extreme outliers in terms of profitability and the use of ROE (rather than operating profitability) present less of a problem in the preceding sections in which only profitability *rankings* are required. As explained below, the results presented in those sections are not sensitive to alternative measures of profitability.

TABLE 8

Relative Firm Profitability and Earnings Sensitivity to Industry News

Panel A: Observations Excluding Operating Profitability Below −1 and Above +1 and ΔE in the Top and Bottom 10 Percent

Variable	Combined News (A)	Positive Industry News (B)	Negative Industry News (C)
Constant	0.004***	0.005***	0.002
	(8.40)	(7.03)	(1.51)
ΔIE	0.481***	0.463***	0.387***
	(23.33)	(16.94)	(4.90)
rp_{t-1}	−0.003***	−0.001	0.001
	(−3.02)	(−0.47)	(0.41)
$\Delta IE * rp_{t-1}$	0.259***	0.144***	0.701***
	(6.78)	(2.83)	(4.48)
Obs.	31,988	25,970	6,018
Adj. R^2	0.122	0.063	0.065

Panel B: Observations Excluding Operating Profitability Below −0.75 and Above +0.75 and ΔE in the Top and Bottom 10 Percent

Variable	Combined News (A)	Positive Industry News (B)	Negative Industry News (C)
Constant	0.004***	0.006***	0.002
	(8.59)	(7.72)	(1.63)
ΔIE	0.468***	0.438***	0.402***
	(22.80)	(16.17)	(5.10)
Rp_{t-1}	−0.003***	−0.001	0.000
	(−3.18)	(−0.58)	(0.07)
$\Delta IE * rp_{t-1}$	0.262***	0.147***	0.672***
	(6.87)	(2.89)	(4.37)
Obs.	31,419	25,521	5,898
Adj. R^2	0.121	0.060	0.065

*** Indicates significance at the 0.01 level.

This table reports the results of the following regression:

$$\Delta E^i_{k,t} = \alpha + \beta_1 \Delta IE^i_t + \beta_2 rp^i_{k,t-1} + \beta_3 \Delta IE^i_{k,t} * rp^i_{k,t-1} + \varepsilon^i_{k,t} \quad (11)$$

where $\Delta E^i_{k,t}$ is the change in operating income after depreciation in year t relative to year t−1 of firm k belonging to industry i (scaled by the year t beginning operating assets); ΔIE^i_t is the change in industry i's operating income after depreciation in period t relative to period t−1 (scaled by the year t beginning operating assets of industry i), with industry operating income equal to the total operating income of all firms in the industry, and $rp^i_{k,t-1}$ is firm k's relative operating profitability (equal to operating income after depreciation divided by beginning operating assets) in industry i in period t−1. Positive industry news is proxied by a positive change in industry earnings, and negative industry news by a negative change in industry earnings. The sample in Panel A is restricted to operating profitability between −1 and +1, excluding the top and bottom 10 percent of $\Delta E^i_{k,t}$. The sample in Panel B is restricted to operating profitability between −0.75 and +0.75, excluding the top, and bottom 10 percent of $\Delta E^i_{k,t}$. We use the OLS procedure, and correct standard errors for heteroscedasticity and cross and serial correlations of the residuals. The numbers in parentheses are t-statistics.

separate tests. We also employ two alternative size proxies: (1) the market value of equity and (2) total assets. The results for both size groups are consistent with our hypotheses, regardless of the size proxy used.

Second, we add relative size (*rs*) and its interaction with industry returns ($IR * rs$) in Regression (7) as additional controls. We define *rs* as the ranking of a firm's size in its industry in year $t-1$ (in ascending order starting from 1) divided by the number of firms in the industry in period $t-1$. Again, we use the market value of equity and the book value of total assets as alternative size proxies. We find that our main variable of interest, $IR * rp$, continues to exhibit the predicted qualitative effects. At the same time, $IR * rs$ has a negative effect in positive industry news situations and a positive effect in negative industry news situations, which is in contrast to the predicted effect of relative profitability. These further tests show that the role of relative profitability is distinct from that of firm size.

Alternative Measures of Relative Profitability

In addition to the ROE and cost efficiency rankings used in the main analysis, we also employ several other definitions of relative profitability: (1) industry median-adjusted ROE; (2) operating profitability rankings; and (3) return on assets rankings. The results are qualitatively the same as those in the main analysis.

Interestingly, when we include industry median-adjusted ROE and its interaction with *IR* as additional variables in (7), we find that $IR * rp$ (as used in the main analysis) continues to exhibit a significant effect (although its magnitude is diminished), which suggests that the effect of the profitability ranking is not totally subsumed by that of industry-adjusted ROE.

Alternative Measures of Industry News

In the primary analysis, we classified the direction of industry news based on the sign of industry returns. Here, we adopt two other ways of classifying positive versus negative industry news for sensitivity checks: (1) industry returns adjusted for market returns, and (2) changes in industry earnings. The results remain qualitatively the same.

To reduce the influence of extreme firm returns on the industry average return, we winsorize firm returns at the top and bottom 5 percent (and, alternatively, 10 percent) in an industry-year. To address the concern that some industry events cause simply a redistribution of profits within an industry (creating winners and losers) as opposed to a general shift in the industry condition that affects firms similarly, we also conduct analyses that exclude those industry-years for which (1) the industry average return is not significantly different from zero at the 5 percent (and, alternatively, 10 percent) level and/or (2) at least 30 percent (alternatively, 20 percent, or 10 percent) of the firms have returns in the opposite direction to that of the overall industry return. We find similar results.

We also use value-weighted industry returns (excluding the firm being examined) as an alternative to equal-weighted returns and obtain similar results.

Indirect Measure of Capacity Utilization

In Section V, we examined cross-industry variation in the asymmetry of the role played by relative profitability in relation to positive and negative industry news by differentiating industries based on capital intensity. In additional analysis, we also examine cross-industry variation by separating industries based on future capital investment, intended as an (indirect) proxy for capacity utilization. We measure an industry's future investment by the total investment in years +1 and +2, scaled by PPE. We predict that the asymmetry in the role of relative profitability (as posited in H2) should be more prominent in the group of high, versus low, future investment (suggesting a higher level of current capacity utilization).[17] Our results are consistent with this prediction.

Additional Robustness Checks

We also perform separate analyses for young and mature firms, for industries with different degrees of concentration, and for the sub-periods of 1973–1988 and 1989–2004. The qualitative results obtained in the main analysis continue to hold. We also conduct analyses concentrating on single-segment firms, using observations with profitability rankings between 25 percent and 75 percent in the industry, using three-digit SICs (instead of four-digit SICs) to identify industries, and using the Fama and MacBeth (1973) test procedure. We also use the two-digit SIC industry classification, which substantially increases the number of firms in an industry-year, and find results similar to those reported above (including the results of Regression (10), shown in Table 7, concerning the co-movement of relative firm profitability and industry beta over time).

X. SUMMARY AND CONCLUDING REMARKS

This study provides theory and empirical evidence to demonstrate that relative profitability plays a role in influencing a firm's return sensitivity to industry-level news. We extend the standard industrial organization (IO) models to examine how firms within the same industry adjust their operational decisions in response to common exogenous shocks, thus leading to changes in firm value. We predict that the stock returns of the less profitable firms in a given industry are more sensitive to industry-wide news than those of their more profitable counterparts, and that this cross-firm difference in return sensitivity is more pronounced in situations of positive rather than negative industry news, especially in industries with a high level of capital intensity.

Consistent with these predictions, we find that the returns of the less profitable firms in an industry co-move with industry returns to a greater extent than those of more profitable firms. Moreover, the variation in this relative profitability-induced return sensitivity to industry returns across firms is greater in the subsample of positive industry returns than in that of negative industry returns, with the asymmetry between the two types of news more pronounced in high (versus low) capital intensity industries. Our results remain the same after controlling for stock price movements that are attributable to firm-specific accounting variables and risk proxies, are robust to alternative definitions of profitability and industry news, and hold in different time periods of the sample and for different firm size and industry concentration subsamples. Finally, we demonstrate that these results are not caused spuriously by the inclusion of multiple-segment firms in the sample.

This study contributes to both academic research and investment management practice in several ways. First, it is among the first to explore the role of relative profitability in determining a firm's return sensitivity to industry returns. In this regard, it extends the literature by showing how the firm-specific information provided in financial reporting interacts with aggregate-level information to explain returns.

Second, looking to how systematic risk is defined in asset-pricing models, we may view a firm's return sensitivity to industry returns as a form of systematic risk faced by investors. By demonstrating how the return sensitivities of firms in the same industry are driven by their competitive positions (as indicated by relative profitability), this study sheds light on the link between a firm's systematic risk and its business fundamentals.

---

[17] In practice, how much firms actually invest is also affected by factors separate from genuine production needs, such as the inability to raise capital and managerial self-interest (e.g., empire building), so the quality of our proxy for capacity utilization is an open empirical issue.

Finally, the findings presented here are applicable to investment management. As less profitable firms are more sensitive to changes in industry-wide conditions, investors with specific insights about the future development of an industry (whose view is not yet reflected in market prices) should pick stocks within the industry, conditional on their outlook. That is, they should shift their investment allocation toward less profitable firms in the industry when they expect the overall conditions of the industry to improve (to capture more capital appreciation) and, conversely, shift to more profitable firms when they expect the industry's conditions to deteriorate (to better preserve capital).

REFERENCES

Arai, M. 2003. Wages, profits, and capital intensity: Evidence from matched worker-firm data. *Journal of Labor Economics* 21 (3): 593–618.

Baldwin, R. 1971. Determinants of the commodity structure of U.S. trade. *American Economic Review* 61 (1): 126–146.

Beaver, W., P. Kettler, and M. Scholes. 1970. The association between market-determined and accounting-determined risk measures. *The Accounting Review* 45 (4): 654–682.

Beaver, W., and J. Manegold. 1975. The association between market-determined and accounting-determined measures of systematic risk: Some further evidence. *Journal of Financial and Quantitative Analysis* 10 (2): 231–284.

Biddle, G., P. Chen, and G. Zhang. 2001. When capital follows profitability: Non-linear residual income dynamics. *Review of Accounting Studies* 6 (2–3): 229–265.

Burgstahler, D., and I. Dichev. 1997. Earnings, adaptation, and equity value. *The Accounting Review* 72 (2): 187–215.

Chan, K. 1988. On the contrarian investment strategy. *The Journal of Business* 61 (2): 147–163.

Chen, P., and G. Zhang. 2007. How do accounting variables explain stock price movements? Theory and evidence. *Journal of Accounting and Economics* 43 (2–3): 219–244.

De Bondt, W., and R. Thaler. 1989. Anomalies: A mean-reverting walk down Wall Street. *Journal of Economic Perspectives* 3 (1): 189–202.

Dechow, P., A. Hutton, and R. Sloan. 1999. An empirical assessment of the residual income valuation model. *Journal of Accounting and Economics* 26 (1–3): 1–34.

Dechow, P., and R. Sloan. 1997. Returns to contrarian investment strategies: Tests of naive expectations hypotheses. *Journal of Financial Economics* 43 (1): 3–27.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin. 2003. Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research* 41 (5): 797–836.

Fama, E., and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (2): 427–465.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.

Fama, E., and K. French. 1996. Multifactor explanation of asset pricing anomalies. *Journal of Finance* 51 (1): 55–84.

Fama, E., and K. French. 2008. Dissecting anomalies. *Journal of Finance* 63 (4): 1653–1678.

Fama, E., and J. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–636.

Foster, G. 1981. Intra-industry information transfers associated with earnings releases. *Journal of Accounting and Economics* 3 (3): 201–232.

Han, J., and J. Wild. 1990. Unexpected earnings and intra-industry information transfer: Further evidence. *Journal of Accounting Research* 28 (1): 211–219.

King, B. 1966. Market and industry factors in stock price behavior. *Journal of Business* 39 (1): 139–190.

Lakonishok, J., A. Shleifer, and R. Vishny. 1994. Contrarian investment, extrapolation, and risk. *Journal of Finance* 49 (5): 1541–1578.

Leontief, W. 1954. Domestic production and foreign trade: The American capital position re-examined. *Economia Internationale* 7 (1): 3–32. Reprinted in *Readings in International Economics*, edited by R. E. Caves and H. G. Johnson, 503–527. The American Economic Association. Homewood, Ill.: Richard D. Irwin.

Magee, R. 1974. Industry-wide commonalities in earnings. *Journal of Accounting Research* 12 (2): 270–287.

Ramash, S. 2002. Investor and analyst reactions to earnings announcements of related firms: An empirical analysis. *Journal of Accounting Research* 40 (5): 1351–1376.

Roll, R. 1988. R^2. *Journal of Finance* 43 (3): 541–566.

Rosenberg, B., and W. McKibben. 1973. The prediction of systematic and specific risk in common stocks. *Journal of Financial and Quantitative Analysis* 8 (2): 317–333.

Ryan, S. 1997. A survey of research relating accounting numbers to systematic equity risk, with implications for risk disclosure policy and future research. *Accounting Horizons* 11 (2): 82–95.

Ryan, S., and P. Zarowin. 2003. Why has the contemporaneous linear returns-earnings relation declined? *The Accounting Review* 78 (2): 523–553.

Thomas, J., and F. Zhang. 2008. Overreaction to intra-industry information transfers? *Journal of Accounting Research* 46 (4): 909–940.

Winston, G. 1979. On measuring factor productions in industries with different seasonal and shift patterns. Or did the Leontief paradox ever exist? *The Economic Journal* 89 (356): 897–904.

Zhang, G. 2000. Accounting information, capital investment decisions, and equity valuation: Theory and empirical implications. *Journal of Accounting Research* 38 (2): 271–295.

APPENDIX A

EXTENSIONS OF THEORETICAL ANALYSIS

To extend the theoretical analysis in Section II, we consider a number of additional scenarios: differences in cost efficiency driven by fixed costs; shocks on the slope (as opposed to the intercept) of the demand function; and Bertrand (price) competition with differentiated products.

Differences in Cost Efficiency Resulting from Fixed Costs

Assume that the cost of firm i ($i = 1,2$) to produce quantity x_j is $F_j + Cx_j$, where F_j is the fixed cost of production for firm j and C is the variable cost per unit of output (identical for the two firms). Let firm 1 be more cost-efficient than firm 2; that is, $F_1 < F_2$. The firms engage in Cournot competition, and they determine output quantities to maximize their respective profits:

$$\underset{x_j}{Max}[a - b(x_1 + x_2) - C]x_j - F_j, j = 1,2. \tag{A1}$$

Following the approach in Section II, we derive the equilibrium earnings as $E_j = (a - C)^2/(9b) - F_j$, $j = 1,2$. The firm values are:

$$V_j = (a - C)^2/(9br_j) - F_j/r_j, \ j = 1,2. \tag{A2}$$

A demand shock (Δa) prompts the firms to adjust their production, thus changing their equilibrium earnings. Firm values after the shock become $V_j' = (a + \Delta a - C)^2/(9br_j) - F_j/r_j, j = 1,2$, and therefore the returns triggered by the demand shock are:

$$R_j \equiv \frac{V_j' - V_j}{V_j} = \frac{2(a - C)\Delta a}{(a - C)^2 - 9bF_j}, \ j = 1,2. \tag{A3}$$

Given $F_1 < F_2$, we have $R_2 > R_1 > 0$ if $\Delta a > 0$ and $R_2 < R_1 < 0$ if $\Delta a < 0$. Hence, as in Section II, the return of the less profitable firm (firm 2) is more sensitive to the demand shock.

Shocks on the Slope of the Demand Curve

Consider the same basic problem as the one above in which the cost functions consist of both a fixed and a variable component. Firm value prior to a demand shock is given by (A2). Now, a shock occurs unexpectedly on the slope of the demand function, denoted by Δb, which prompts the firms to revise their quantities. Firm values after the shock are $V_j' = (a - C)^2/[9(b + \Delta b)r_j] - F_j/r_j, j = 1,2$. Therefore, the returns triggered by the shock are:

$$R_j \equiv \frac{V_j' - V_j}{V_j} = \frac{-(a - C)^2 \Delta b}{(a - C)^2(b + \Delta b) - 9b(b + \Delta b)F_j}, \; j = 1, 2. \tag{A4}$$

Again, given $F_1 < F_2$, the return of firm 2 (the less profitable firm) is more sensitive to the industry shock (Δb).

Bertrand Competition with Product Differentiation

Now assume that the firms engage in Bertrand (price) competition with differentiated products (if nothing else, the location of a supplier can be a differentiating factor). The respective demand functions for the firms' products are $X_1 = a + P_2 - P_1$, and $X_2 = a + P_1 - P_2$. For simplicity, we ignore fixed costs. Let $C_1 < C_2$, such that firm 1 has a lower variable cost per unit of output than firm 2. The firms choose selling prices to maximize profit; that is:

$$\underset{P_1}{Max}(P_1 - C_1)(a + P_2 - P_1), \quad \text{and}$$

$$\underset{P_2}{Max}(P_2 - C_2)(a + P_1 - P_2).$$

The first-order conditions are $a + P_2 - 2P_1 + C_1 = 0$, and $a + P_1 - 2P_2 + C_2 = 0$. The resulting equilibrium prices are $P_1 = a + (2C_1 + C_2)/3$, and $P_2 = a + (2C_2 + C_1)/3$. And the equilibrium earnings are $E_j = [3a + (C_1 + C_2) - 2C_j]^2/9, j = 1,2$.

As before, we set firm value as the capitalization of earnings. A common demand shock, Δa, triggers all firms to revise their prices (and quantities), thus resulting in a change in firm value. The returns that arise from the demand shock are $R_1 = 6\Delta a/(3a + C_2 - C_1)$, and $R_2 = 6\Delta a/(3a + C_1 - C_2)$. Again, given $C_1 < C_2$, the return of firm 2 (the less profitable firm) is more sensitive to common shocks.

CALCUTTA UNIVERSITY

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1349–1381

American Accounting Association
DOI: 10.2308/accr-10035

The Pricing of Accruals Quality: January versus the Rest of the Year

Christina A. Mashruwala
Shamin D. Mashruwala
Baruch College–CUNY

ABSTRACT: As an alternative way to shed light on the debate over whether accruals quality is a priced risk factor, we examine the effect of seasonality on the pricing of the modified Dechow and Dichev (2002) accruals quality measure (AQ). We find that (1) high AQ stocks outperform low AQ stocks *only* in January; (2) during the rest of the calendar year, high AQ firms underperform low AQ firms such that there is no AQ premium on an annual basis; (3) about half of the January AQ premium occurs during the first five trading days of January; (4) a January AQ premium is observed in almost every year of our sample period; and (5) the January AQ premium reflects, at least partly, the stock price effects of tax loss selling around the turn of the year. Taken together, these findings are difficult to reconcile with a risk interpretation of accruals quality.

Keywords: *accruals quality; systematic risk; January effect; seasonality; tax loss selling.*

I. INTRODUCTION

Whether information risk independently affects a firm's cost of equity capital continues to be a controversial issue. In the accounting literature, an important but unresolved question is whether accruals quality (AQ), an aspect of information risk and often proxied by the modified Dechow and Dichev (2002) AQ measure, is a priced risk factor. Some argue in favor of a risk interpretation (e.g., Francis et al. 2005; Aboody et al. 2005; Ogneva 2008; Kim and Qi 2010), while others reject this view (e.g., Core et al. 2008; hereafter CGV). Despite this ambiguity about AQ's risk status, a large literature continues to make inferences based on the assumption that AQ proxies for information risk (Appendix A lists a selection of recently published articles that view AQ as a risk factor). A better understanding of whether AQ's pricing is compatible with a risk interpretation therefore seems important.

We thank Steven Kachelmeier and Wayne Thomas (the editors), two anonymous reviewers, Donal Byard, and Carol Marquardt for helpful comments and suggestions.

Editor's note: Accepted by Wayne Thomas.

Submitted: November 2008
Accepted: December 2010
Published Online: April 2011

A fundamental issue in this debate is whether there is a robust, positive relation between AQ and future returns.[1] If high AQ firms do not outperform low AQ firms on average, then AQ cannot represent risk and additional asset-pricing tests, as in CGV, are moot. However, prior evidence on the relation between AQ and future returns is inconclusive, differing across studies, sample periods, and methodologies. For example, CGV conclude that there is no robust positive relation between AQ and future returns after controlling for market microstructure biases and/or transaction costs. On the other hand, Ogneva (2008) and Kim and Qi (2010) argue that high AQ firms do outperform low AQ firms after controlling for cash flow shocks or removing low-priced stocks, respectively.

Since traditional asset-pricing tests (as in the above studies) have yielded mixed results on AQ's pricing, and thereby on the risk debate, we take a different approach and examine how seasonality affects AQ's pricing. In particular, we investigate the role of January versus the rest of the calendar year. As we show, viewing the pricing of AQ through a different lens—that of seasonality—is a powerful way of shedding light on whether AQ reflects risk. The seasonality approach also helps to explain why prior evidence on AQ's pricing (based on tests using all calendar-year months) has so far been ambiguous.

Our exploration of January's role in AQ's pricing is motivated by two considerations. First, the nature of the AQ measure strongly suggests that it could proxy for tax loss selling around the turn of the year. If so, then January is likely to play an important role in AQ's pricing because tax loss selling causes (at least partly) the well-known January effect in stock prices (Roll 1983b; Reinganum 1983; Sias and Starks 1997; Poterba and Weisbenner 2001).[2] Prior research argues that firms with high operating or return volatility—such as small firms—are more likely candidates for tax loss selling because they are statistically more likely to experience operating losses, and therefore negative stock returns, in any given year (e.g., Roll 1983b; Constantinides 1984). This reasoning suggests that AQ could also proxy for tax loss selling because high AQ firms exhibit many of the characteristics that increase the likelihood of negative stock returns in any given year: higher stock return volatility, higher sales and earnings volatility, a higher frequency of operating losses, poorer financial health, weaker internal controls, a greater likelihood of GAAP violations and accounting restatements, and a higher incidence of negative earnings surprises and/or negative cash flow shocks (Dechow and Dichev 2002; Francis et al. 2005; Doyle et al. 2007; Ogneva 2008; Kim and Qi 2010).

Second, a finding that the AQ premium is concentrated in January—which seems plausible if AQ proxies for tax loss selling—would cast serious doubt on the notion that AQ reflects risk.[3] This is because there does not appear to be any theoretical rationale for systematic risk to manifest itself only in January. Indeed, the fact that other hypothesized risk factors exhibit a premium primarily, and in some cases only, in January—such as beta (Tinic and West 1984), size (Blume and Stambaugh 1983), book-to-market (Loughran 1997), and idiosyncratic volatility (Doran et al. 2010)—has often been viewed as being inconsistent with a risk interpretation of these variables (e.g., Roll 1983b; Tinic and West 1984; Loughran 1997).[4]

---

[1] Higher AQ implies poorer accruals quality. Note that, throughout this study, AQ refers specifically to the modified Dechow and Dichev (2002) AQ measure; it does not refer to accruals quality in general.

[2] The tax loss selling argument for the January effect runs as follows. To lock in capital losses for tax purposes, investors sell loser stocks at year-end and buy them back in January. The resulting price pressure leads to positive abnormal returns in January. We discuss tax loss selling and its possible link to AQ in Section II.

[3] In this study, the term "AQ premium" does not imply a risk premium to AQ. It simply refers to a positive relation between AQ, considered as a firm characteristic, and future abnormal returns. We explicitly use the term "AQ risk premium" to refer to a risk premium to AQ.

[4] As we discuss in Section IV, the calendar-year pricing of AQ differs from the pricing of size and book-to-market in two fundamental ways. These differences indicate that the pricing of AQ is even less suggestive of risk than is the pricing of size and book-to-market.

Using the modified Dechow and Dichev (2002) AQ measure, we find that the abnormal return to an AQ hedge portfolio is positive *only* in the month of January. During our sample period, the average abnormal January return to an AQ hedge portfolio (based on equal-weighted deciles) is an economically and statistically significant 6.94 percent. Equally surprising is that about half of this January premium, on average, is driven by the first five trading days in January. However, for the rest of the year, taken as a whole or in any individual month, there is no evidence of an AQ premium. In fact, after controlling for firm size, we observe an AQ *discount* in the non-January period. That is, the January AQ premium reverses over the rest of the year, often in such a way that there is no premium on an annual basis. The premium is also extremely stable, occurring in 35 of the 38 sample years. Finally, we find that the magnitude of the January AQ premium is increasing in tax loss selling, consistent with the premium being driven, at least partly, by the stock price effects of tax loss selling around the turn of the year.

The January AQ effect is not subsumed by the size, beta, book-to-market, idiosyncratic volatility, dividend-yield (Keim 1986), and reverse momentum (Jegadeesh and Titman 1993) January effects, is not confined to low-priced stocks (Bhardwaj and Brooks 1992), is not driven by market microstructure effects such as illiquidity, and is robust to controlling for cash flow shocks (Ogneva 2008). It also cannot be explained by the "information hypothesis" (Rozeff and Kinney 1976), according to which the January effect reflects heightened information uncertainty—and perhaps greater risk—toward the end of the calendar year and the subsequent resolution of such uncertainty in January. Inferences are also similar using value-weighting, monthly or daily returns, and different sample periods. Finally, we find similar results for two other commonly used proxies for accruals quality.

Viewing the pricing of AQ from the perspective of seasonality contributes to the literature in at least three ways. First, it sheds a deeper light on the risk debate by revealing that January plays a critical role in AQ's pricing. An AQ premium is seen only in January and about half of the January premium occurs during the first trading week in January. Conversely, the non-January period exhibits an AQ discount—i.e., high AQ stocks underperform low AQ stocks during this period—such that there is typically no AQ premium on an annual basis. These findings are difficult to reconcile with the view that AQ is a systematic risk factor; that is, it is unclear why such risk would be confined only to January, why much of the risk in January would manifest itself during the first trading week, and why a January risk premium would reverse completely during the rest of the year. Rather, these findings, together with the fact that tax loss selling drives (at least partly) the January AQ premium, strongly suggest that the premium reflects a temporary mispricing in January that is subsequently corrected.

Second, our findings help reconcile the conflicting prior evidence on the fundamental issue of whether high AQ firms earn higher returns on an annual basis. Our results suggest that the ambiguity in prior research arises not from the use of different methodologies and/or market microstructure biases, but rather due to the role of January. High AQ firms outperform low AQ firms only in January, and it is the outsized January premium that occasionally results in an annual premium being observed.[5] When January is excluded, high AQ firms never earn higher returns on an annual basis, regardless of methodology or sample period.

---

[5] For instance, we find no annual return premium to high AQ stocks during January 1971–November 2003 (used by CGV) and April 1971–March 2002 (used by Francis et al. 2005), but we dc find such a premium during January 1985–November 2003 (used by Aboody et al. 2005). This inconsistency is due to the temporal distribution of the AQ effect during the calendar year—although the size of the January AQ premium is similar in the three periods, its non-January reversal is weaker in the January 1985–November 2003 period, resulting in the observed annual AQ premium for this period. When January is excluded, there is no annual AQ premium for any of these sample periods.

Finally, our results extend the January effect literature, most of which investigates the size effect in January. We show that another firm characteristic, accruals quality, also exhibits a powerful and independent January effect. In addition, like the small-firm January effect, the January AQ effect also seems to be driven, at least partly, by tax loss selling.

In Section II, we discuss prior literature on the pricing of AQ, the role of January in asset pricing, the relation between the January effect and tax loss selling, and the plausible link between AQ and tax loss selling. In Sections III, IV, and V, we discuss the methodology for our main portfolio tests, the results, and robustness checks, respectively. Section VI concludes.

II. PRIOR RESEARCH

The Pricing of AQ

In the theoretical literature, there seems to be no consensus yet about whether accounting quality/information risk should be priced (Easley and O'Hara 2004; Hughes et al. 2007; Lambert et al. 2007; Armstrong et al. 2009; Christensen et al. 2010). The empirical literature on this issue is equally inconclusive. In the interests of brevity, we confine our discussion of this literature to studies most relevant in our context—those that examine the pricing of the modified Dechow and Dichev (2002) AQ measure, a widely used proxy for information risk, using realized stock returns rather than estimates of implied cost of capital.

Francis et al. (2005) conclude that AQ is a priced risk factor. Their inference is partially based on a positive factor loading on an AQ factor-mimicking portfolio in Fama and French (1993) time-series regressions (after controlling for standard risk factors). However, as CGV point out, a positive factor loading in Fama and French (1993) regressions does not imply that the AQ factor bears a risk premium and is therefore insufficient to conclude that AQ is a priced risk factor.

To address this criticism, several studies examine two issues: whether returns of high AQ firms outperform those of low AQ firms and, if so, whether the outperformance of high AQ firms is due to risk. The overall evidence for both these issues is mixed and appears to depend on the specific sample period and/or methodology used. The reason(s) for such conflicting findings is not entirely clear.

Aboody et al. (2005) find no relation between AQ and future returns, at least for a subset of insiders' buy-and-sell transactions (they do not test this relation for their full sample). They also find a statistically insignificant risk premium for an earnings-quality factor for all four of their earnings quality measures (including AQ), which is inconsistent with earnings quality being a priced risk factor. However, they attribute their weak results to "noise contained in the observations, primarily from the [high earnings quality] firms" (2005, 665), and conclude that their overall evidence is still consistent with earnings quality being a priced risk factor.

Ecker et al. (2006) do not directly test whether AQ predicts future returns. However, their AQ factor-mimicking portfolio (using daily returns) earns a positive daily return on average, indicating that AQ and future returns are highly positively correlated during their sample period. They do not test whether this outperformance of high AQ firms is due to risk.

Using monthly returns, CGV find a positive relation between AQ and future returns in the period tested by Aboody et al. (2005), but no relation in the period tested by Francis et al. (2005).[6] In contrast, using daily returns and daily/monthly rebalancing of equal-weighted portfolios, CGV find a strong positive relation between AQ and future returns for all sample periods examined. When testing for a risk premium to AQ, CGV find a risk premium *only* for the Aboody et al. (2005) period, and not for the other periods they examine. They argue, however, that the inconsistency in their results is due to bid-ask bias and/or transaction costs: upon rebalancing their portfolios annually instead of daily/monthly, or dropping stocks with a price less than $5, they find that AQ neither predicts future returns nor bears a risk premium. CGV (2008, 14) acknowledge, however, that annual rebalancing is "not completely satisfying" because less frequent rebalancing (annual instead of monthly) may involve a loss of information about the pricing of AQ and bias their tests toward the null.

---

[6] CGV do not directly test whether AQ predicts future returns for the Aboody et al. (2005) period. However, the fact that they find a risk premium to AQ for this period implies a positive relation between AQ and future returns.

In contrast to CGV, two more recent studies find a positive relation between AQ and future returns as well as a positive risk premium to AQ after controlling for cash flow shocks (Ogneva 2008) and low-priced stocks (Kim and Qi 2010). Ogneva (2008) argues that controlling for cash flow shocks—which affect realized returns—is important. Without such a control, the estimated effect of AQ on future realized returns could be biased downward because AQ is highly correlated with negative cash flow shocks. Kim and Qi (2010) repeat CGV's tests, but find a positive risk premium to AQ only when stocks less than $5 are excluded from their sample. Their findings appear to contradict those of CGV, who find some evidence of an AQ risk premium only when stocks less than $5 are *included* in the sample.

Despite such mixed evidence on AQ's risk status, the notion that information risk (proxied by AQ) is priced remains widespread. A large literature that examines the role of information risk in different settings continues to rely on AQ as a proxy. Even studies that do not use AQ as a risk proxy frequently rely on Francis et al.'s (2004, 2005) conclusion that information risk is priced. Appendix A briefly describes recently published articles in top-tier accounting journals that bear on this topic. The fact that many researchers continue to view AQ as a proxy for information risk underscores the importance of resolving the ambiguity about AQ's pricing and its suitability as a proxy for information risk.

The Role of January in Testing Hypothesized Risk Factors

An important motivation for investigating the role of January in AQ's pricing is that, historically, January has played a crucial role in casting doubt on whether certain firm characteristics—such as size, beta, and book-to-market—are priced risk factors. A finding that January plays a similar role in AQ's pricing could, in the same way, shed a deeper light on whether AQ proxies for systematic information risk.

Although both size and book-to-market exhibit an abnormal return premium on an annual basis, this annual premium is highly (but not entirely) concentrated in January. Keim (1983) and Blume and Stambaugh (1983), among others, show that more than half of the annual size premium is concentrated in January, while more than half of the January size premium occurs during the first trading week in January. Similarly, for the book-to-market effect, Davis (1994) finds that the annual return premium is entirely due to January for the 1940–1963 period. For the more recent 1963–1995 period, Loughran (1997) finds that for all but the two smallest size quintiles—representing about 94 percent of total market capitalization—the book-to-market premium is captured entirely in the month of January.

More surprisingly, even theoretical risk models such as the CAPM, the APT, and Merton's (1987) incomplete-information CAPM seem to describe returns only in January. Tinic and West (1984), among others, show that high CAPM-beta stocks earn higher returns only in January, while Gultekin and Gultekin (1987) show that, when January returns are excluded, there is no statistically significant relation between expected returns and the risk factors predicted by the APT model. Similarly, Doran et al. (2010) find that idiosyncratic volatility (a proxy for idiosyncratic risk) is positively related to future returns—as predicted by Merton (1987) if idiosyncratic risk is priced—only in January.[7]

---

[7] This finding may not generalize to all measures of idiosyncratic risk used in the literature. See Doran et al. (2010) for a discussion of this issue.

Although the reason(s) why return premiums to these firm characteristics are concentrated largely, and sometimes entirely, in January continues to be debated, researchers have often viewed such a calendar-year return pattern as being inconsistent with a risk explanation for several reasons (e.g., Roll 1983b; Tinic and West 1984; Loughran 1997). First, existing asset-pricing models do not predict any special role for the month of January, let alone explain why risk would be priced *only* in January. It is possible that risk varies during the year, for example, due to seasonal macroeconomic fluctuations (Ogden 2003). However, such time-varying risk does not necessarily imply any unique role for January. Even if January were somehow riskier from the point of view of seasonality, it is hard to see why seasonal risk would be confined only to January (as for beta and idiosyncratic volatility) and why it would be so heavily concentrated during the first five trading days of January (as for size).

Second, there is little empirical evidence to support a risk explanation for the January effect. Ritter and Chopra (1989) and Kramer (1994) reject seasonal variation in CAPM-betas or risk premia as an explanation for the small-firm January effect. However, Kramer (1994) argues that a multifactor APT model (based on macroeconomic factors) with seasonal risk and risk premia can account for the small-firm January effect. His reasoning, however, does not explain why January rather than some other month plays such a unique role, or why seasonal risk would manifest itself exclusively in January (as for beta and idiosyncratic volatility). Several studies also test the "information hypothesis," first proposed by Rozeff and Kinney (1976), as a plausible risk explanation for the January effect. According to this hypothesis, since the majority of firms have December fiscal year-ends, the January effect reflects the greater uncertainty—and therefore the greater systematic risk—of these firms around the turn of the year. The greater uncertainty of these firms arises from the fact that they are expected to release important new financial information in January. These firms therefore earn higher returns in January to compensate for their higher risk. The overall evidence, however, is inconsistent with this hypothesis (Brauer and Chang 1990; Reinganum and Gangopadhyay 1991; Kim 2006).[8]

Third, a large literature shows that the January effect is caused at least partially by year-end tax loss selling by individual investors and/or by year-end "window dressing" by institutional money managers, consistent with the January effect being caused by mispricing rather than risk. We discuss the tax loss selling explanation in the next section.[9]

We show that January also plays a critical role in AQ's pricing. Not only does AQ predict higher returns only in January, but also the complete reversal of this premium during the rest of the calendar year, along with the fact that the January premium at least partly reflects mispricing due to tax loss selling, is difficult to reconcile with the view that AQ proxies for systematic information risk. We discuss this issue more fully in presenting our results (Section IV).

---

[8] Kim (2006) argues that the increase in information risk in January stems from the greater information uncertainty surrounding firms' annual financial results. However, he finds the January effect for both December and non-December fiscal year-end firms, which appears to be inconsistent with his hypothesis. His argument also does not explain why more than half of the small-firm January effect occurs during the first trading week in January; given that most December fiscal year-end firms do not disclose any new information during the first five trading days of January, it is hard to see how information uncertainty is resolved (and why prices rise) during this period. Moreover, since a majority of firms also have calendar quarter-ends and release important new financial information thereafter, the information risk story would also predict a January-like return effect around calendar quarter-ends. However, there is no evidence of systematically higher returns for small firms in these months.

[9] Tax loss selling and window dressing are the two most commonly cited reasons for the January effect. According to the latter, because portfolio managers are evaluated in relation to their peers, they have an incentive to sell loser stocks and buy winner stocks around the turn of the year in order to make their portfolio holdings look better. The evidence on this hypothesis is less extensive and less compelling than that for tax loss selling (e.g., Athanassakos 1992; Griffiths and White 1993; Ng and Wang 2004). Moreover, studies attempting to disentangle the tax loss selling and window dressing hypotheses tend to support the tax loss selling explanation (e.g., Sias and Starks 1997; Poterba and Weisbenner 2001; Grinblatt and Moskowitz 2004).

Tax Loss Selling and the January Effect

The most widely investigated explanation for the January effect is tax loss selling at year-end by individual investors. The intuition for this argument is as follows. By realizing capital losses before year-end, individual investors can reduce their tax liability by using the realized losses to offset either their realized capital gains or their adjusted gross income. In addition, because the short-term capital gains tax rate is higher than the long-term rate (except during 1988–1990), investors can maximize the ability of capital losses to shelter taxable income by realizing the losses while they are still short-term. On the other hand, investors have little incentive to realize capital gains because deferring gains reduces their tax burden. In particular, deferring short-term capital gains would qualify them to be taxed at the lower long-term capital gains tax rate when the gains are eventually realized. Thus, given the tax benefit of locking in capital losses, taxable investors sell their loser stocks by year-end and buy them back in January. In the absence of sufficient rational arbitrage, the stock price pressure resulting from such tax loss selling leads to positive abnormal returns in January for these loser stocks.[10]

A sizable literature documents evidence consistent with tax loss selling at least partly driving the small-firm January premium (e.g., Reinganum 1983; Roll 1983b; Sias and Starks 1997; Poterba and Weisbenner 2001). However, tax loss selling does not seem to completely explain the premium. For example, countries with no capital gains taxes (e.g., Japan), and/or countries where the tax year does not end in December (e.g., the U.K. and Australia), also exhibit the January effect (Gultekin and Gultekin 1983; Berges et al. 1984; Kato and Schallheim 1985).

AQ and Tax Loss Selling

Based on the evidence that tax loss selling causes a January effect, we hypothesize that January could play an important role in AQ's pricing due to the possibility that AQ proxies for tax loss selling. Roll (1983b) and Constantinides (1984) argue that since firms with greater stock return or operating volatility are statistically more likely to have experienced capital gains/losses in any given year, they would be more likely candidates for tax loss selling, all else equal. Brauer and Chang (1990) and Brickley et al. (1991), among others, report that firms with greater return variance do indeed exhibit greater tax loss selling. Roll's (1983b) finding that the small-firm January effect is partly due to tax loss selling is also consistent with more volatile firms—such as small firms—being more likely candidates for tax loss selling.

The above reasoning implies that AQ, like size, may also proxy for tax loss selling because AQ happens to be very highly correlated with stock return volatility (Kim and Qi 2010) as well as operating volatility and the frequency of operating losses (Dechow and Dichev 2002; Francis et al. 2005).[11] Moreover, high AQ firms exhibit poorer financial health, weaker internal controls, a greater likelihood of GAAP violations and accounting restatements, and a higher incidence of negative earnings surprises and/or negative cash flow shocks (e.g., Doyle et al. 2007; Ogneva 2008; Kim and Qi 2010). For all these reasons, high AQ firms are statistically more likely to experience negative stock returns, and therefore capital losses for tax purposes, in any given year.

---

[10] In theory, the optimal tax strategy can be quite complex and does not necessarily imply tax loss selling only at year-end. Constantinides (1984) shows that when short-term capital gains tax rates are higher than long-term rates and there are transaction costs, it can be optimal for investors to realize losses just before they become long-term, or just before the year ends, whichever is earlier. He predicts that tax loss selling gradually increases from January to December and then dries up at the beginning of January.

[11] The correlations between AQ and earnings volatility, accruals volatility, operating cash flow volatility, and the frequency of losses are 0.82, 0.75, 0.60, and 0.63, respectively (Dechow and Dichev 2002).

Consistent with AQ proxying for tax loss selling, we find that (1) the pricing of AQ displays a strong and independent January effect; (2) an AQ premium is observed only in January; and (3) the size of the January AQ premium varies positively with other proxies for tax loss selling. This evidence suggests that the AQ premium in January is caused (at least partly) by mispricing attributable to tax loss selling around the turn of the year.

III. DATA, VARIABLES, AND METHODOLOGY

Data

For our portfolio tests, the full sample period runs from January 1971 through March 2008.[12] The sample includes all common stocks listed on the NYSE, AMEX, and NASDAQ during our sample period, subject to data requirements for the calculation of our variables. Monthly and daily stock returns and prices are obtained from CRSP, annual accounting data are obtained from Compustat, and the Fama and French (1993) factors, the risk-free rate, and the momentum factor are obtained from Kenneth French's website. We obtain institutional ownership data from the CDA/Spectrum institutional money manager holdings (13-F) database, while earnings announcement dates are from I/B/E/S. Our illiquidity measure, based on Amihud (2002), is obtained from Joel Hasbrouck's website.[13] We describe below the measurement of our main variables, *AQ* and *SIZE*. All other variables are defined in Appendix B.

Measurement of *AQ* and *SIZE*

We compute *AQ* following Francis et al. (2005). Their measure of *AQ* is based on the Dechow and Dichev (2002) model, augmented with the change in revenue and the level of property, plant, and equipment (*PPE*). We begin by estimating a regression of total current accruals (*TCA*) on lagged, current, and future cash flows (*CFO*) plus the change in revenue and the level of *PPE*. All variables are deflated by average total assets.

$$TCA_{j,t} = a_{0j} + a_{1j}CFO_{j,t-1} + a_{2j}CFO_{j,t} + a_{3j}CFO_{j,t+1} + a_{4j}\Delta Rev_{j,t} + a_{5j}PPE_{j,t} + \varepsilon_{j,t} \qquad (1)$$

where:

$TCA = (\Delta CA - \Delta Cash) - (\Delta CL - \Delta STDEBT)$;
$TA = TCA - Dep$;
ΔCA = change in current assets;
$\Delta Cash$ = change in cash/cash equivalents;
ΔCL = change in current liabilities;
$\Delta STDEBT$ = change in short-term debt;
Dep = depreciation and amortization expense;
$NIBE$ = net income before extraordinary items;
$CFO = NIBE - TA$;
ΔRev = change in revenue; and
PPE = gross property, plant, and equipment.

We estimate Equation (1) in the cross-section each year for each industry with at least 20 observations in a given year. Industries are defined using Fama and French's (1997) 48-industry classification. Estimation of Equation (1) generates firm- and year-specific residuals. For each firm

---

[12] We replicated our main tests for three other sample periods used by prior research: January 1971–November 2003 (used by CGV), April 1971–March 2002 (used by Francis et al. 2005), and January 1985–November 2003 (used by Aboody et al. 2005). The inferences are similar to those reported for the full sample period.

[13] We thank Kenneth French and Joel Hasbrouck for providing the data.

j, we calculate accruals quality (*AQ*) for year *t* as the standard deviation of firm-specific residuals from year *t*–4 through year *t* (i.e., the five-year standard deviation). Accruals quality is assumed to be poorer when the standard deviation of the residuals (*AQ*) is higher. For each firm, *AQ* is measured at every fiscal year-end and is used to form portfolios starting three months after the fiscal year-end. For example, we use the value of *AQ* for year *t* calculated for a firm with a December fiscal year-end to form portfolios starting in April of year *t*+1. The next section describes the portfolio formation procedure.

We compute *SIZE* monthly for each firm as the market capitalization of the firm at the end of the month prior to portfolio formation. Market capitalization is defined as the share price times the number of common shares outstanding on the last trading day of the month. For example, *SIZE* calculated at the end of March of year *t* is used to form portfolios in April of year *t*. Although *AQ* is calculated annually, we compute *SIZE* monthly in order to better control for size and isolate the effect of *AQ* on stock returns.[14]

Portfolio Formation

In any given month, we form portfolios on *AQ* and/or *SIZE* using all firms for which we have returns in that particular month. For example, based on the measurement of *AQ* and *SIZE* described above, if we sort firms on *AQ* and *SIZE* in April 2000, we effectively sort on *AQ* values calculated between January 1999 and December 1999 (depending on a firm's fiscal year-end), and we sort on *SIZE* measured as of the end of March 2000.

Firms are sorted into deciles when forming portfolios on *AQ* alone. When forming portfolios on both *AQ* and *SIZE*, we sort firms into quintiles, resulting in a total of 25 *AQ* × *SIZE* portfolios each month. All reported results are based on independent portfolio sorts on *AQ* and *SIZE*; results using conditional sorts are similar.

Computation of Abnormal Returns

After forming portfolios every month as described above, we compute monthly (daily) equal-weighted and value-weighted portfolio returns. Thus, for each portfolio, we obtain 447 (9,399) monthly (daily) returns over our full sample period from January 1971 to March 2008. To compute the average monthly (daily) abnormal return for each portfolio during the sample period, we estimate four-factor Fama and French (1993) time-series regressions for each portfolio using the 447 (9,399) monthly (daily) portfolio returns. That is, for each portfolio, we estimate the following model:

$$R_{pt} - R_{ft} = b_{0p} + b_{1p}(R_{mt} - R_{ft}) + b_{2p}SMB_t + b_{3p}HML_t + b_{4p}UMD_t + \varepsilon_{p,t} \qquad (2)$$

where R_p represents the equal-weighted or value-weighted monthly (daily) return for the portfolio and R_f is the monthly (daily) yield of the one-month Treasury bill. The four contemporaneous independent variables are the monthly (daily) excess return on the CRSP value-weighted market portfolio ($R_m - R_f$), the difference in monthly (daily) returns to value-weighted portfolios containing small- and large-firm stocks (*SMB*), the difference in monthly (daily) returns to value-weighted portfolios containing high and low book-to-market stocks (*HML*), and the difference in monthly (daily) returns to value-weighted portfolios containing firms with high and low prior momentum (*UMD*).[15] The intercept b_0 represents the average monthly (daily) abnormal return to the relevant portfolio.

---

[14] Computing *SIZE* annually instead of monthly does not affect our inferences.

[15] Inferences are similar if the momentum factor (*UMD*) is excluded.

We also estimate Equation (2) for AQ hedge portfolios that are long in the highest AQ portfolio and short in the lowest AQ portfolio, using the monthly (daily) hedge portfolio return as the dependent variable. In these regressions, our primary variable of interest is the estimated intercept because it can be interpreted as the average monthly (daily) abnormal return generated by the AQ hedge portfolio over the sample period.

Controls

Since the bid-ask bias in low-priced stocks is likely to be of particular concern around the turn of the year (Blume and Stambaugh 1983; Keim 1989), we try to control for it in three different ways. First, we report value-weighted portfolio returns. Value-weighting considerably mitigates the bid-ask bias in a non-arbitrary manner (Roll 1983a; Blume and Stambaugh 1983). However, a possible disadvantage is that value-weighting likely understates the AQ effect because the latter is much stronger among smaller stocks, perhaps for reasons that have nothing to do with bid-ask bias. Second, when using daily returns, we report equal-weighted portfolio returns after removing stocks priced lower than $1 for which the bias is likely to be most severe (Ball et al. 1995). Finally, we explicitly control for both stock price and illiquidity in our cross-sectional regression tests to isolate this bias.[16]

In our cross-sectional tests, we also control for the beta, dividend yield, book-to-market, momentum, and idiosyncratic volatility January effects. Incorporating these controls in a cross-sectional framework avoids the practical problem of inadequately small portfolio sizes when portfolios are formed on more than three variables. Table 1 shows descriptive statistics and correlations for the main variables used in this study.

IV. RESULTS

The AQ Effect: January versus Non-January

We begin by documenting how AQ spreads stock returns on an annual basis and then examine the effect of seasonality on the predictive ability of AQ. The first two columns of Table 2 show equal- and value-weighted monthly abnormal returns to an AQ hedge portfolio based on AQ deciles (*AQ10 – AQ1*). For simplicity, we do not tabulate the estimated coefficients for the three Fama and French (1993) factors and momentum. We evaluate statistical significance using Newey-West standard errors with two lags to correct for serial correlation in monthly portfolio returns. All reported significance tests are two-tailed.

There is no evidence that AQ predicts future returns when we use all calendar months in the sample period. The average equal- and value-weighted abnormal returns to the AQ hedge portfolio are 0.43 percent and –0.01 percent per month, respectively, and both are statistically insignificant at conventional levels ($p = 0.13$ and 0.96, respectively).

Next, we examine the AQ hedge portfolio abnormal return for January and all non-January months, both individually and taken together. Now we observe an AQ premium, but it is entirely a January phenomenon and seems to reverse completely over the rest of the calendar year such that AQ does not spread annual returns. In January, the equal-weighted (value-weighted) abnormal return is an economically large 6.94 percent (2.44 percent) and statistically significant ($p < 0.01$ and $p = 0.04$,

---

[16] Dropping stocks less than $5 instead leads to similar inferences. In our tests, excluding low-priced stocks reduces the magnitude of portfolio returns. This is consistent with an overstatement of portfolio returns due to bid-ask bias among low-priced stocks, a finding documented by many prior studies (e.g., Blume and Stambaugh 1983) as well as by CGV. Unlike these studies, Kim and Qi (2010) find that the return premium to AQ increases (rather than decreases) when stocks less than $5 are excluded from the sample.

TABLE 1

Descriptive Statistics

Panel A: Descriptive Statistics

	Mean	Std. Dev.	Q1	Median	Q3
AQ	0.046	0.038	0.02	0.035	0.059
BETA	1.111	0.683	0.649	1.049	1.475
SIZE	1185.8	3657.92	24.86	111.21	597.65
BM	1.075	2.157	0.333	0.618	1.08
MOM	0.167	0.453	−0.09	0.14	0.385
YIELD	0.025	0.069	0.00	0.003	0.028
IVOL	0.02	0.026	0.006	0.011	0.023
PRICE	19.232	17.8	5.875	14.125	27.04
ILLIQ	5.45	35.657	0.014	0.165	1.654
ARET	−0.13	0.544	−0.328	−0.047	0.199
PTLS	−0.749	0.203	−0.917	−0.795	−0.625
IO	0.355	0.277	0.102	0.313	0.573

Panel B: Pearson (top) and Spearman (bottom) Correlations

	AQ	*BETA*	*SIZE*	*BM*	*MOM*	*YIELD*	*IVOL*	*PRICE*	*ILLIQ*	*ARET*	*PTLS*	*IO*
AQ		0.17	−0.13	−0.07	0.03	−0.15	0.45	−0.31	0.09	−0.01	0.28	−0.22
BETA	0.16		−0.05	−0.06	0.03	−0.13	0.26	−0.12	−0.03	−0.02	0.21	0.08
SIZE	−0.28	−0.01		−0.09	0.02	−0.01	−0.15	0.47	−0.05	−0.04	−0.18	0.30
BM	−0.17	−0.09	−0.43		−0.13	0.71	−0.04	−0.04	0.11	0.14	0.13	−0.02
MOM	−0.01	0.01	0.12	−0.29		−0.06	0.17	0.17	0.07	−0.90	−0.49	−0.03
YIELD	−0.48	−0.24	0.22	0.26	−0.05		−0.17	0.13	−0.07	0.07	−0.07	0.06
IVOL	0.53	0.31	−0.45	−0.08	0.05	−0.63		−0.36	0.21	−0.10	0.39	−0.28
PRICE	−0.42	−0.12	0.76	−0.30	0.23	0.42	−0.63		−0.14	−0.21	−0.48	0.50
ILLIQ	0.23	−0.04	−0.93	0.39	0.02	−0.19	0.38	−0.69		0.04	0.14	−0.15
ARET	0.07	0.04	−0.19	0.32	−0.92	0.01	0.07	−0.34	0.12		0.57	−0.02
PTLS	0.26	0.21	−0.40	0.25	−0.53	−0.24	0.45	−0.57	0.31	0.70		−0.24
IO	−0.22	0.08	0.73	−0.21	**−0.01**	0.12	−0.33	0.58	−0.74	−0.09	−0.24	

All variables are winsorized at the 1 percent and 99 percent levels. The sample period runs from January 1971 to March 2008. In Panel B, all correlations are significant at the 1 percent level, except the one in bold. All variables are defined in Appendix B.

respectively). However, for the non-January months taken together, the average equal-weighted (value-weighted) abnormal return is −0.25 percent (−0.33 percent) per month but insignificantly different from zero at conventional levels ($p = 0.15$ in both cases). (Later, we show that the statistical insignificance of the non-January AQ discount is due to the confounding effect of firm size. Once we control for size, the non-January reversal of the January AQ premium is clearly visible and statistically significant.) More tellingly, the abnormal return is significantly positive only in January. In every other calendar month, the abnormal return is either zero or negative.[17]

---

[17] The positive and significant value-weighted abnormal return in May seems to be a statistical fluke. Not only is the equal-weighted abnormal return in May insignificantly different from zero, but also both the equal- and value-weighted abnormal returns in May are insignificantly different from zero in every other sample period we examined.

TABLE 2

AQ Hedge Portfolio Abnormal Returns and *AQFactor* Returns (Using Monthly Returns)

	AQ Hedge Portfolio Abnormal Returns (*AQ10* − *AQ1*)		
Month	**Equal-Weighted**	**Value-Weighted**	***AQFactor* Returns**
January	6.94***	2.44**	4.44***
	(4.87)	(2.09)	(5.24)
February	0.76	0.59	0.73
	(1.52)	(0.74)	(1.37)
March	−0.32	−1.07	−0.27
	(−0.89)	(−1.46)	(−0.67)
April	−0.44	−0.44	−0.56
	(−0.89)	(−0.87)	(−1.29)
May	0.24	1.24**	0.30
	(0.56)	(2.04)	(0.72)
June	−0.75*	−0.23	−0.02
	(−1.96)	(−0.44)	(−0.06)
July	−0.08	−0.36	−0.42
	(−0.19)	(−0.38)	(−1.58)
August	−0.85**	0.12	−0.50
	(−2.47)	(0.17)	(−1.41)
September	0.34	−1.32	−0.19
	(0.68)	(−1.47)	(−0.55)
October	0.13	−0.39	−0.91
	(0.26)	(−0.68)	(−1.61)
November	0.00	−0.29	−0.10
	(0.00)	(−0.42)	(−0.21)
December	−1.03	0.24	−0.94**
	(−1.25)	(0.37)	(−2.43)
All Months	0.43	−0.01	0.14
(n = 447)	(1.53)	(−0.05)	(0.82)
Non-January Months	−0.25	−0.33	−0.26
(n = 409)	(−1.44)	(−1.44)	(−1.51)

---

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed.

This table presents the average monthly equal-weighted and value-weighted abnormal returns to a hedge portfolio formed on AQ (*AQ10* − *AQ1*) and the average monthly return for an AQ factor-mimicking portfolio (*AQFactor*). To calculate the hedge portfolio returns, AQ deciles are formed each month. The average monthly hedge portfolio abnormal return is the intercept in a time-series regression of the monthly hedge portfolio return on the monthly Fama and French (1993) factors ($R_m - R_f$, *SMB*, *HML*) and the monthly momentum factor (*UMD*). To calculate the *AQFactor*, the sample is sorted into AQ quintiles at the beginning of each month. The *AQFactor* return for the month is the difference in equal-weighted returns between the top two quintiles and the bottom two quintiles. Returns are presented in percentages. t-statistics, shown in parentheses below the return, are calculated using Newey-West standard errors with two lags. The sample period runs from January 1971 to March 2008. All variables are defined in Appendix B.

To test whether AQ is a risk factor, Francis et al. (2005) construct a factor-mimicking portfolio based on AQ (*AQFactor*).[18] The third column in Table 2 shows that this factor-mimicking portfolio also displays a strong January effect. Moreover, the *AQFactor* return is significantly positive only in January (return = 4.44 percent, $p < 0.01$). For all calendar months taken together, for all non-January months taken together, and for any individual non-January month, the *AQFactor* return is never significantly positive.

Figure 1, Panel A shows that the equal-weighted AQ January effect is not driven solely by the two extreme AQ deciles. The returns are monotonically increasing across deciles in January. For all other calendar months, returns across AQ deciles are more or less similar and there is no upward trend. For the value-weighted AQ decile returns in January (Figure 1, Panel B), the increase across deciles is more gradual and non-monotonic, highlighting the importance of firm size for the January AQ premium.

Is the January AQ Premium Compensation for Risk?

Given the apparent lack of any theoretical or empirical justification for systematic risk to be priced primarily in January, the fact that high AQ firms earn higher returns *only* in January strongly suggests that AQ may not reflect systematic risk. We now discuss several additional findings that, when considered together with the results above, reinforce the impression that AQ is unlikely to be proxying for risk.

First, we find that the January AQ premium is extremely stable from year to year. The equal-weighted January return on an AQ hedge portfolio (based on deciles) is positive in 35 of the 38 sample years, with a high average return of 8.42 percent (Table 3) and a relatively low standard deviation of 8.2 percent (not tabulated). Such a persistently positive return to a trading strategy is more indicative of mispricing than of risk. This is because, if the average January premium (over time) earned by the AQ hedge portfolio were compensation for bearing systematic risk, then one would expect the portfolio to periodically experience losses over a 38-year period. Otherwise, it is not clear in what sense such a portfolio is risky (see, for example, Lakonishok et al. [1994] and Bernard et al. [1997] for further discussion of this issue).

Second, we find that about half of the January AQ effect occurs during the first trading week of January. Table 3 shows equal- and value-weighted returns to an AQ hedge portfolio for January, the first trading day in January, and the first five trading days in January. For each firm in AQ deciles 1 and 10, we compute buy-and-hold returns over the five-day period. We then average the five-day buy-and-hold return across all firms in the relevant portfolio. The five-day AQ hedge portfolio return is the difference between the five-day return for decile 10 and the five-day return for decile 1. The value-weighted portfolio returns are based on firms' market capitalization as of the last trading day of the preceding December.

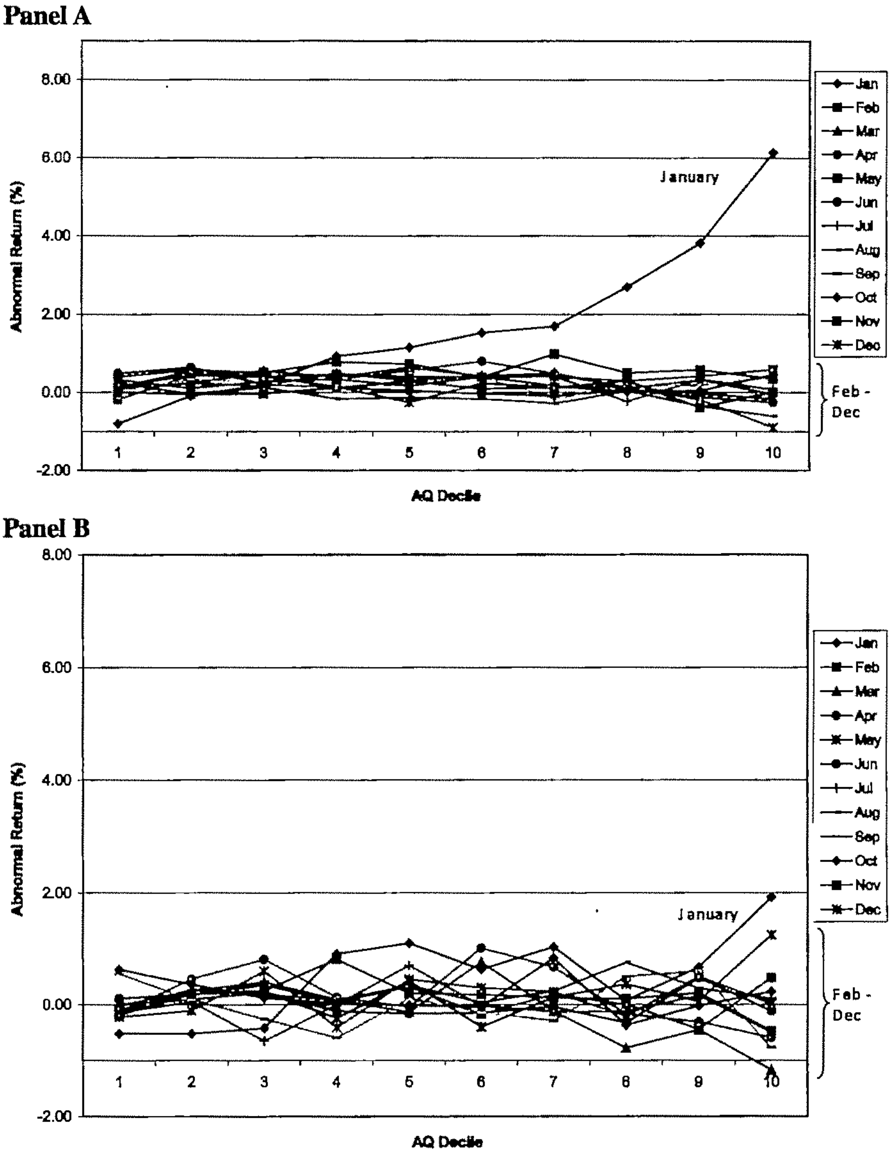
Using either means or medians, about 50 percent of the January return to the equal-weighted AQ hedge portfolio occurs during the first trading week in January, while about 17 percent of it occurs on the first trading *day* of the year. Such a lopsided concentration of the AQ premium in early January is difficult to reconcile with a risk explanation. Even if the January premium were compensation for some kind of seasonal risk that manifests itself only in January (although there does not appear to be any justification in the literature for such a notion of risk), one would expect the premium to be spread uniformly throughout January—i.e., the proportion of the premium

---

[18] The *AQFactor* is constructed as follows. At the beginning of each month, the sample is sorted into AQ quintiles. The *AQFactor* return for the month is the difference in equal-weighted returns between the top two quintiles and the bottom two quintiles (see Francis et al. [2005] for further details).

FIGURE 1
AQ Effect by Calendar Month Panel

Panel A: Equal-Weighted Portfolio Abnormal Returns
Panel B: Value-Weighted Portfolio Abnormal Returns



This figure plots equal-weighted (Panel A) and value-weighted (Panel B) abnormal returns for each AQ decile by calendar month. Returns are presented in percentages. The abnormal return for each portfolio is the intercept in a time-series regression of monthly portfolio excess returns on the monthly Fama and French (1993) factors ($R_m - R_f$, *SMB*, *HML*) and the monthly momentum factor (*UMD*). The sample period runs from January 1971 to March 2008. All variables are defined in Appendix B.

TABLE 3

Timing of the AQ Effect in January

		AQ Hedge Portfolio Returns (*AQ10 − AQ1*)				
		January	**One-Day**	**% of January**	**Five-Day**	**% of January**
Equal-Weighted	Mean	8.42***	1.41***	16.75%	4.11***	48.81%
		(4.86)	(6.91)		(7.74)	
	Median	6.65	1.12	16.84%	3.66	55.04%
Value-Weighted	Mean	4.07***	0.32	7.86%	1.08*	26.54%
		(3.48)	(1.46)		(1.94)	
	Median	1.59	0.28	17.61%	1.77	111.32%

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed.

This table presents the mean and median raw returns to a hedge portfolio formed on AQ (*AQ10 − AQ1*) for the entire month of January, the first trading day in January, and the first five trading days in January. AQ deciles are formed each month. To calculate the hedge portfolio return for the first five trading days in January, we compute the five-day buy-and-hold return for each firm, and then compute the equal- or value-weighted hedge portfolio return for the five-day period. Returns are presented in percentages. t-statistics, shown in parentheses below the return, are calculated using Newey-West standard errors with two lags. The sample period runs from January 1971 to March 2008. All variables are defined in Appendix B.

earned in the first five trading days would only be about 23.8 percent rather than the observed 50 percent (assuming an average of 21 trading days per month).

Third, we find that the January AQ premium reverses itself completely over the rest of the calendar year, such that there is no AQ premium on an annual basis. This again suggests that the January AQ premium is due to a temporary mispricing that subsequently corrects itself. If the premium were due to some kind of January risk, then one would not expect it to subsequently reverse.

Fourth, we find that the January AQ premium is increasing in tax loss selling, consistent with the January AQ premium being caused (at least partly) by mispricing related to tax loss selling around the turn of the year (discussed later in this section).

Fifth, we find evidence inconsistent with the January AQ premium being due to risk arising from information uncertainty around the turn of the year (see Section V).

Finally, note that the calendar-year pricing of AQ is less consistent with a risk explanation than is the pricing of either size (e.g., Keim 1983) or book-to-market (e.g., Loughran 1997), in two fundamental ways. Unlike the AQ premium, the size and book-to-market premiums occur in January as well as in the non-January period—i.e., they are not confined only to January, even though a large proportion of the annual premium occurs in January. Also unlike AQ, both size and book-to-market exhibit economically large premiums on an annual basis—i.e., the size and book-to-market January premiums do not reverse over the rest of the calendar year.

The above findings, especially when taken together, are difficult to reconcile with a risk interpretation of AQ, given the current theoretical and empirical understanding of risk in the literature. Rather, they suggest that the January AQ premium reflects mispricing, at least partly due to tax loss selling. We caution, however, that our findings should not be taken as conclusive evidence against the risk argument. Although the tests we use are standard in the asset-pricing literature, the equilibrium returns model they assume could be misspecified if researchers' current understanding of risk is somehow incomplete or flawed. At a minimum, however, any potential risk explanation for AQ must now satisfactorily explain how risk is compatible with the collective evidence presented here.

Firm Size and the January AQ Effect

In Table 4, we examine how firm size influences the relation between AQ and future stock returns. In addition to clarifying how the AQ effect differs between small and large firms, partitioning on size also controls for the January size anomaly. The results show that the January AQ premium is independent of firm size and occurs in four out of five size quintiles. Conversely, for the non-January period, not only is there is no AQ premium for any size quintile, but we see a significant AQ *discount* in three out of five size quintiles.[19]

As described earlier, we independently sort all firms into size and AQ quintiles each month. For each size quintile, Table 4 reports the equal- and value-weighted abnormal return to an AQ hedge portfolio based on the two extreme quintiles (*AQ5* – *AQ1*). For simplicity, we only discuss the results for equal-weighted portfolios, as the value-weighted results are generally similar. In Panel A, for all calendar months taken together, the average monthly AQ abnormal hedge return is significantly positive only for the smallest size quintile (return = 0.53 percent, $p = 0.05$). However, this annual AQ effect is entirely driven by January. When January is excluded, the AQ abnormal return becomes insignificantly different from zero (return = 0.04 percent, $p = 0.87$).

The January AQ effect remains robust after controlling for size, and is positive and statistically significant in four of the five size quintiles. For the smallest size quintile, the January AQ premium is a large 6.09 percent ($p < 0.01$), which decreases to 0.76 percent ($p = 0.19$) for the largest size quintile. Thus, although the January AQ effect is considerably stronger for smaller stocks, it is not driven by the January size anomaly.

For the non-January period, we now observe a statistically significant AQ discount in three of the five size quintiles—i.e., for these size quintiles, AQ is negatively related to future returns. Given the generally insignificant abnormal returns to AQ on an annual basis, this finding suggests that the January AQ effect reverses itself completely during the rest of the calendar year. Recall that this reversal of the January AQ premium is statistically insignificant in the absence of a size control (Table 2). Controlling for size is therefore important for a clearer understanding of how the pricing of AQ varies during the year.

Using Daily Instead of Monthly Returns

We repeated the above tests using daily returns because some prior studies (e.g., CGV) also use daily returns to test whether AQ is priced. The results, shown in Table 5, are similar to those using monthly returns. For the sake of brevity, we do not discuss these results.

Cross-Sectional Regressions

Our results for the January AQ premium are robust to various additional controls. In Table 6, we employ monthly Fama and MacBeth (1973) cross-sectional regressions to simultaneously control for several firm characteristics that are known to be associated with January returns: firm size (*SIZE*; e.g., Keim 1983), the CAPM beta (*BETA*; e.g., Tinic and West 1984), book-to-market (*BM*; e.g., Loughran 1997), dividend yield (*YIELD*; e.g., Keim 1986), idiosyncratic volatility (*IVOL*; e.g., Doran et al. 2010), stock price (*PRICE*; e.g., Bhardwaj and Brooks 1992), and momentum (*MOM*; e.g., Jegadeesh and Titman 1993). We also add a proxy for illiquidity (*ILLIQ*; based on Amihud [2002]) as an additional control for market microstructure effects and/or liquidity risk. *AQ* and

---

[19] Bhardwaj and Brooks (1992) argue that the small-firm January premium is a stock price phenomenon rather than a size phenomenon, i.e., size plays no role in the January premium after controlling for stock price. We emphasize the role of size in our tests to be consistent with the literature on the January effect, which continues to focus on size rather than stock price. We do, however, control for stock price in both our cross-sectional tests and in unreported portfolio tests, and find that the January AQ premium is not subsumed by stock price.

TABLE 4

AQ × *SIZE* Hedge Portfolio Abnormal Returns (Using Monthly Returns)

Panel A: Equal-Weighted Portfolio Abnormal Returns

	All Months (n = 447)			January (n = 38)			Non-January Months (n = 409)		
SIZE Quintile	*AQ1*	*AQ5*	Hedge (*AQ5* − *AQ1*)	*AQ1*	*AQ5*	Hedge (*AQ5* − *AQ1*)	*AQ1*	*AQ5*	Hedge (*AQ5* − *AQ1*)
1	0.86***	1.39***	0.53**	3.43***	9.52***	6.09***	0.51**	0.55***	0.04
	(3.88)	(4.24)	(1.97)	(3.66)	(5.95)	(4.34)	(2.49)	(2.58)	(0.16)
2	0.24*	−0.05	−0.29	1.70***	4.12***	2.42**	0.07	−0.44***	−0.51***
	(1.82)	(−0.27)	(−1.53)	(3.29)	(3.63)	(2.18)	(0.58)	(−2.93)	(−2.77)
3	0.18*	−0.18	−0.37*	−0.93***	1.33**	2.26***	0.28***	−0.41***	−0.68***
	(1.93)	(−1.14)	(−1.94)	(4.06)	(2.22)	(3.45)	(2.79)	(−3.15)	(−4.40)
4	0.12	−0.08	−0.20	−1.38***	0.84**	2.22***	0.24***	−0.19	−0.43***
	(1.59)	(−0.59)	(−1.36)	(−3.85)	(2.31)	(4.53)	(3.06)	(−1.45)	(−3.01)
5	0.01	0.06	0.05	−1.05***	−0.30	0.76	0.12*	0.13	0.01
	(0.16)	(0.45)	(0.36)	(−2.77)	(−0.50)	(1.34)	(1.74)	(0.97)	(0.05)

Panel B: Value-Weighted Portfolio Abnormal Returns

	All Months (n = 447)			January (n = 38)			Non-January Months (n = 409)		
SIZE Quintile	*AQ1*	*AQ5*	Hedge (*AQ5* − *AQ1*)	*AQ1*	*AQ5*	Hedge (*AQ5* − *AQ1*)	*AQ1*	*AQ5*	Hedge (*AQ5* − *AQ1*)
1	1.05***	1.16***	0.11	3.17***	8.04***	4.87***	0.73***	0.43**	−0.30
	(5.12)	(3.94)	(0.42)	(4.04)	(6.27)	(4.35)	(3.86)	(2.26)	(−1.37)
2	0.70***	0.40**	−0.30	1.94***	4.30***	2.36**	0.55***	0.02	−0.53***
	(5.48)	(2.13)	(−1.52)	(3.96)	(3.89)	(2.11)	(4.74)	(0.10)	(−2.83)
3	0.66***	0.35**	−0.31	−0.62**	1.34**	1.96***	0.77***	0.17	−0.60***
	(6.66)	(2.27)	(−1.64)	(−2.71)	(2.49)	(3.28)	(7.35)	(1.30)	(−3.64)

(continued on next page)

TABLE 4 (continued)

SIZE Quintile	All Months (n = 447) *AQ1*	All Months (n = 447) *AQ5*	All Months (n = 447) Hedge (*AQ5* – *AQ1*)	January (n = 38) *AQ1*	January (n = 38) *AQ5*	January (n = 38) Hedge (*AQ5* – *AQ1*)	Non-January Months (n = 409) *AQ1*	Non-January Months (n = 409) *AQ5*	Non-January Months (n = 409) Hedge (*AQ5* – *AQ1*)
4	0.60***	0.43***	–0.17	–0.97**	1.39***	2.37***	0.72***	0.33**	–0.39***
	(7.32)	(3.13)	(–1.12)	(–2.49)	(3.74)	(4.38)	(8.85)	(2.44)	(–2.68)
5	0.51***	0.57***	0.06	0.02	1.44*	1.42	0.57***	0.49***	–0.07
	(9.47)	(3.45)	(0.34)	(0.06)	(1.73)	(1.47)	(10.30)	(2.80)	(–0.39)

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed.

This table shows, for each size quintile, the average monthly abnormal return for the extreme AQ quintiles (*AQ1* and *AQ5*), and the AQ hedge portfolio (*AQ5* – *AQ1*). AQ and *SIZE* quintiles are formed monthly using independent sorts. The average monthly portfolio abnormal return is the intercept in a time-series regression of the monthly portfolio excess return on the monthly Fama and French (1993) factors ($R_m - R_f$, *SMB*, *HML*) and the monthly momentum factor (*UMD*). Returns are presented in percentages. t-statistics, shown in parentheses below the return, are calculated using Newey-West standard errors with two lags. The sample period runs from January 1971 to March 2008. All variables are defined in Appendix B.

TABLE 5

AQ and *AQ* × *SIZE* Hedge Portfolio Abnormal Returns (Using Daily Returns)

Panel A: AQ Hedge Portfolios

AQ Hedge Portfolio Abnormal Returns
(*AQ10* – *AQ1*)

Equal-Weighted			Value-Weighted		
All (n = 9,399)	**January (n = 797)**	**Non-January (n = 8,602)**	**All (n = 9,399)**	**January (n = 797)**	**Non-January (n = 8,602)**
0.024***	0.252***	0.00	0.016	0.161***	0
(3.53)	(8.91)	(−0.07)	(1.62)	(4.11)	(0.05)

Panel B: *AQ* × *SIZE* Hedge Portfolios

AQ Hedge Portfolio Abnormal Returns
(*AQ5* – *AQ1*)

	Equal-Weighted			Value-Weighted		
***SIZE* Quintile**	**All (n = 9,399)**	**January (n = 797)**	**Non-January (n = 8,602)**	**All (n = 9,399)**	**January (n = 797)**	**Non-January (n = 8,602)**
1	0.009	0.164***	−0.006	0.028***	0.177***	0.012
	(0.93)	(4.09)	(−0.61)	(2.85)	(4.17)	(1.20)
2	−0.011*	0.080***	−0.019***	−0.008	0.087***	−0.017**
	(−1.69)	(3.01)	(−2.93)	(−1.20)	(3.09)	(−2.44)
3	−0.020***	0.078***	−0.030***	−0.018***	0.071***	−0.027***
	(−3.10)	(3.10)	(−4.55)	(−2.74)	(2.83)	(−4.02)
4	−0.015**	0.085***	−0.025***	−0.013**	0.089***	−0.023***
	(−2.33)	(3.60)	(−3.76)	(−1.98)	(3.60)	(−3.36)
5	0.007	0.054*	0.002	0.013	0.098***	0.005
	(0.98)	(1.88)	(0.32)	(1.54)	(2.88)	(0.55)

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed.

This table shows the average daily abnormal returns for a hedge portfolio formed on AQ (Panel A) or AQ and *SIZE* (Panel B). In Panel A, AQ deciles are formed each month. In Panel B, AQ and *SIZE* quintiles are formed monthly using independent sorts. The average daily hedge portfolio abnormal return is the intercept in a time-series regression of the daily hedge portfolio return on the daily Fama and French (1993) factors ($R_m - R_f$, *SMB*, *HML*) and the daily momentum factor (*UMD*). Returns are presented in percentages. To mitigate bid-ask bias in computing equal-weighted portfolio returns, stocks with a price of less than $1 at the beginning of each month are dropped from the sample. t-statistics, shown in parentheses below the return, are calculated using Newey-West standard errors with five lags. The sample period runs from January 1971 to March 2008. All variables are defined in Appendix B.

YIELD are computed at fiscal year-end (and mapped to returns three months later), *ILLIQ* is measured at calendar year-end, and all other variables are computed every month. Each month, firm-specific returns are regressed on firm-specific, lagged independent variables, as follows:

$$R_{it} = b_0 + b_1 AQ_{it-1} + b_2 Controls_{it-1} + \varepsilon_{it}. \tag{3}$$

The monthly coefficient estimates from Equation (3) are then averaged over the sample period to obtain the mean coefficient estimates. Statistical significance of the mean coefficient estimates is based on Newey-West standard errors (with two lags) of the monthly coefficient estimates to adjust for autocorrelation in the latter. All significance tests are two-tailed. Note that our sample period for

TABLE 6

Fama-MacBeth Cross-Sectional Regressions

$$R_{it} = b_0 + b_1 AQ_{it-1} + b_2 Controls_{it-1} + \varepsilon_{it} \qquad (3)$$

	All Months (n = 432)			January (n = 36)			Non-January Months (n = 396)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.014***	0.02***	0.021***	0.031***	0.059***	0.060***	0.012***	0.016***	0.017***
	(6.16)	(7.50)	(8.69)	(3.01)	(8.42)	(8.93)	(5.45)	(6.03)	(7.00)
AQ	0.02	0.014	0.004	0.905***	0.263***	0.276***	−0.06**	0.009	−0.021*
	(0.72)	(1.08)	(0.29)	(6.25)	(4.78)	(5.07)	(−2.49)	(−0.71)	(−1.83)
BETA		0.001	0.004		0.011**	0.014**		0.000	−0.001
		(0.41)	(0.30)		(2.50)	(2.34)		(−0.35)	(−0.66)
ln(SIZE)		−0.001***	−0.001***		−0.007***	−0.007***		−0.001*	−0.001**
		(−3.32)	(−3.84)		(−5.30)	(−6.91)		(−1.72)	(−2.36)
ln(BM)		0.005***	0.003***		0.016***	0.019***		0.004***	0.002***
		(7.04)	(4.49)		(5.28)	(5.80)		(5.79)	(2.62)
MOM		0.012***			−0.015**			0.014***	
		(8.43)			(−2.36)			(10.56)	
YIELD		0.004	0.001		0.125***	0.122***		−0.007	−0.01
		(0.53)	(0.13)		(3.01)	(3.69)		(−0.90)	(−1.43)
IVOL		−0.130***	−0.086**		0.699***	0.616***		−0.205***	−0.15***
		(−3.63)	(−2.52)		(5.19)	(4.92)		(−5.43)	(−4.43)
PRICE		0.001	0.006**		−0.016	−0.024*		0.003	0.009***
		(0.48)	(2.16)		(−1.25)	(−1.86)		(1.10)	(3.17)
ILLIQ		0.015***	0.017***		0.112***	0.112***		0.006	0.009*
		(2.78)	(3.53)		(3.68)	(4.58)		(1.23)	(1.88)
Mean Adj. R^2	0.0106	0.0581	0.0527	0.0264	0.1301	0.1236	0.0092	0.0515	0.0462

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed.

This table presents the time-series mean coefficients from monthly cross-sectional regressions using all monthly returns (Columns 1–3), January returns (Columns 4–6), and non-January returns (Columns 7–9). The dependent variable is monthly returns. All explanatory variables are winsorized at the 1 percent and 99 percent levels; the fiscal year-end explanatory variables are mapped to returns three months later. The coefficients on *PRICE* and *ILLIQ* are multiplied by 100. t-statistics, shown in parentheses below the coefficient estimates, are based on Newey-West standard errors with two lags. The sample period runs from January 1971 to December 2006. All variables are defined in Appendix B.

these tests ends in December 2006 rather than March 2008 because the *ILLIQ* data are available only through 2005. However, our inferences are similar if we use the full sample period and either drop the *ILLIQ* control or use the percentage bid-ask spread to proxy for liquidity.

The results in Table 6 are consistent with the portfolio results. First, when using all calendar months (columns 1–3), the estimated coefficient on *AQ* is never significantly different from zero—i.e., there is no annual AQ premium. Second, the estimated *AQ* coefficient is always strongly positive and highly significant in January (columns 4–6). Third, the estimated coefficient on *AQ* is negative in non-January months (columns 7–9), unless momentum is included in the specification. The reason why momentum subsumes the non-January AQ discount is unclear.[20] Inferences are also similar using decile ranks for the independent variables (untabulated).

Does Tax Loss Selling Explain the January AQ Effect?

Earlier, we hypothesized that AQ could possibly proxy for the incidence of tax loss selling. If tax loss selling drives the January AQ premium, then one would expect (1) the January AQ premium to be greater for firms with greater tax loss selling; and (2) the effect of tax loss selling on the January AQ premium to be greater for firms with lower institutional ownership. This is because tax loss selling is commonly attributed to selling by taxable individual investors rather than institutional investors (e.g., Ritter 1988; Sias and Starks 1997).

To proxy for the likelihood of tax loss selling in a particular firm, we employ two widely used measures in the January effect literature. The first measure is the firm's buy-and-hold return over the previous calendar year, excluding the last five trading days in December (Roll 1983b; Sias and Starks 1997; Poterba and Weisbenner 2001). The idea here is that the lower a firm's return in the prior year, the greater would be the likelihood of tax loss selling in the firm at the end of the year. For ease of interpretation, we multiply the prior year's return by −1 and call this variable *ARET*; thus, the likelihood of tax loss selling is increasing in *ARET*.

The second measure attempts to capture the idea that the greater the short-term capital loss in a particular year, the greater is the probability of tax loss selling. Following Reinganum (1983) and Chang and Pinegar (1986), we measure the likelihood of a short-term capital loss by using the ratio of the stock price on the last trading day of the year (excluding the last five trading days) and the maximum stock price between January 1 and the last trading day of the year (excluding the last five trading days). We define "short-term" as the entire calendar year because the IRS definition of a short-term holding period is 12 months for the majority of our sample period. By construction, this ratio varies between 0 and 1, with a lower value indicating a greater likelihood of tax loss selling. For ease of interpretation, we multiply this ratio by −1 and call the resulting variable *PTLS*; thus, similar to *ARET*, the likelihood of tax loss selling is expected to be increasing in *PTLS*.[21]

---

[20] The pricing of the other firm characteristics is also consistent with prior research; a large proportion of the size or book-to-market premium occurs in January (e.g., Keim 1983; Loughran 1997), beta is priced only in January (e.g., Tinic and West 1984), idiosyncratic volatility is priced positively in January and negatively in the non-January period (Doran et al. 2010), momentum is priced negatively in January and positively in the non-January period (e.g., Jegadeesh and Titman 1993), stock price is negatively correlated with January returns (e.g., Bhardwaj and Brooks 1992), and high-dividend firms earn higher returns in January (Keim 1986).

[21] Note that our *PTLS* measure differs slightly from Reinganum (1983) in two ways. First, he uses a six-month period (July 1–December 31) as his short-term holding period because the IRS holding period was six months for most of his sample period (the holding period was six months from 1971 to 1976 and from July 1984 to 1987, nine months for 1977, and 12 months for the remainder of our sample period). Second, he only excludes the last trading day in December to compute his measure, whereas we exclude the last five trading days. We do this to be consistent with the recent literature on the January effect (e.g., Sias and Starks 1997; Poterba and Weisbenner 2001). However, inferences are similar if we instead exclude only the last trading day in December.

As with our main cross-sectional tests, we use the Fama and MacBeth (1973) methodology for the tax loss selling tests. We estimate the following specification:

$$January\,R_{it} = b_0 + b_1AQ_{it-1} + b_2TLS_{it-1} + b_3AQ_{it-1} \times TLS_{it-1} + b_4Controls_{it-1} + \varepsilon_{it} \quad (4)$$

where *TLS* is the probability of tax loss selling in the previous year (either *ARET* or *PTLS*). If the January AQ premium is caused by the stock price effects of tax loss selling, then the coefficient on *AQ* should vary positively with tax loss selling—i.e., the coefficient on b_3 in Equation (4) should be positive. Given this directional prediction, the significance test for b_3 is one-tailed.

Table 7 shows that the January AQ premium indeed increases with tax loss selling. The *AQ* × *TLS* interaction is positive and statistically significant using both *ARET* (column 2, est. coeff. = 0.22, p = 0.03) and *PTLS* (column 3, est. coeff. = 0.976, p < 0.01) as proxies for tax loss selling. To get a sense for the economic significance of this result, it is helpful to see how the January AQ premium varies with tax loss selling. Using *ARET*, the January AQ premium increases by 63 percent as we move from quartile 1 to quartile 3 of *ARET* (the *AQ* coefficient goes from 0.185 to 0.301).[22] Using *PTLS*, the result is even stronger: the *AQ* coefficient is almost 0 for quartile 1 of *PTLS*—i.e., there is no January AQ premium when tax loss selling is low—and increases to 0.282 for quartile 3 of *PTLS*. In untabulated tests using rank regressions to examine the issue of economic significance (with decile ranks scaled between 0 and 1), the results are also striking: for both *ARET* and *PTLS*, the *AQ* × *TLS* interaction remains positive and significant as before, but the coefficient on *AQ* now becomes insignificant or negative—i.e., there is no January AQ premium when there is no tax loss selling (for decile rank 0). Overall, therefore, tax loss selling appears to play an economically important role in the January AQ premium. However, the fact that an AQ premium is observed in some specifications even when tax loss selling is low (for example, for quartile 1 of *ARET* in column 2 of Table 7) suggests that tax loss selling is not a complete explanation for the January AQ premium.

Next, we examine whether institutional ownership affects the relation between tax loss selling and the January AQ effect. At the end of every December, we form three institutional ownership (*IO*) portfolios, and then estimate Equation (4) separately for each *IO* portfolio. *IO* is defined as the sum of all institutional holdings of the stock scaled by the number of shares outstanding. Since institutional holdings are reported every calendar quarter, we use the most recent holdings for a firm prior to December 1. The number of outstanding shares is measured at the end of the month in which ownership data are reported. As in prior research, we assume that stocks without any reported 13-F institutional holdings have no institutional ownership.

Table 8 shows that, using either *ARET* or *PTLS*, the January AQ premium is increasing in tax loss selling (i.e., the *AQ* × *TLS* interaction is positive) only for firms with low institutional ownership (columns 1 and 4). This is consistent with tax loss selling at least partly causing the January AQ premium, because tax loss selling is likely to be most prevalent in firms with low institutional ownership. However, for the medium and high *IO* portfolios (columns 2–3 and 5–6), there seems to be no relation between tax loss selling and the January AQ effect (the *AQ* × *TLS* interaction is insignificant), despite the fact that these portfolios also exhibit an AQ premium (albeit about half the size of that for the low *IO* portfolio; untabulated). This latter finding again suggests that tax loss selling does not fully explain the AQ premium. Pursuing a comprehensive explanation of the January AQ premium, however, is beyond the scope of this study and we leave this question for future research.

---

[22] The *AQ* coefficient is computed using the estimated regression coefficients and Table 1. For example, for quartile 1 of *ARET* (*ARET* = −0.328 from Table 1), the *AQ* coefficient is computed as 0.257 + 0.22 × −0.328 = 0.185.

TABLE 7

Tax Loss Selling and the January AQ Effect

$$January\ R_{it} = b_0 + b_1 AQ_{it-1} + b_2 TLS_{it-1} + b_3 AQ \times TLS_{it-1} + b_4 Controls_{it-1} + \varepsilon_{it} \qquad (4)$$

		Tax Loss Selling (*TLS*)	
		Proxy is *ARET*	**Proxy is *PTLS***
	(1)	(2)	(3)
Intercept	0.07***	0.068***	0.098***
	(8.21)	(7.98)	(4.35)
AQ	0.276***	0.257***	0.892***
	(4.96)	(4.58)	(3.94)
TLS		0.016	0.032
		(1.52)	(1.51)
AQ × *TLS*		0.22**	0.976***
		(1.93)	(3.44)
BETA	0.01**	0.009**	0.007*
	(2.13)	(2.04)	(1.69)
ln(*SIZE*)	−0.007***	−0.008***	−0.007***
	(−7.10)	(−7.21)	(−7.02)
ln(*BM*)	0.018***	0.017***	0.016***
	(6.08)	(5.94)	(5.84)
MOM	−0.016**	0.01	0.007
	(−2.23)	(1.00)	(1.43)
YIELD	−0.072	−0.069	−0.073
	(−1.48)	(−1.44)	(−1.56)
IVOL	0.726***	0.678***	0.382***
	(5.13)	(5.06)	(3.07)
PRICE	−0.023*	−0.02	−0.014
	(−1.73)	(−1.59)	(−1.26)
ILLIQ	0.114***	0.11***	0.109***
	(4.62)	(4.51)	(4.39)
Mean Adj. R^2	0.1301	0.1342	0.1398

---

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed except for *AQ* × *TLS* (which is one-tailed).

This table presents the time-series mean coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions using January observations only. The dependent variable is monthly returns. All explanatory variables are winsorized at the 1 percent and 99 percent levels; the fiscal year-end explanatory variables are mapped to returns three months later. The coefficients on *PRICE* and *ILLIQ* are multiplied by 100. t-statistics, shown in parentheses below the coefficient estimates, are based on Newey-West standard errors with two lags. The sample period runs from January 1971 to December 2006. All variables are defined in Appendix B.

V. ADDITIONAL TESTS

Information Hypothesis

Table 9 reports evidence that the information hypothesis (discussed in Section II) is unlikely to explain the January AQ effect. First, the magnitude of the January AQ effect is unaffected when December year-end firms (i.e., firms with upcoming annual financial information in January) are dropped from the sample, as the *AQ* coefficient is similar in columns 1 and 2. Second, the magnitude of the January AQ premium is also unaffected if firms with fiscal

TABLE 8

Tax Loss Selling and the January AQ Effect: The Role of Institutional Ownership

$$January\ R_{it} = b_0 + b_1AQ_{it-1} + b_2TLS_{it-1} + b_3AQ \times TLS_{it-1} + b_4Controls_{it-1} + \varepsilon_{it} \qquad (4)$$

	Tax Loss Selling (*TLS*)					
	Proxy is *ARET*			Proxy is *PTLS*		
	(1) *IO* = Low	(2) *IO* = Med	(3) *IO* = High	(4) *IO* = Low	(5) *IO* = Med	(6) *IO* = High
Intercept	0.083***	0.037***	0.006	0.126***	0.067***	0.017
	(8.02)	(4.06)	(0.47)	(4.23)	(3.45)	(0.56)
AQ	0.261***	0.093	0.100*	0.901**	0.158	0.162
	(3.03)	(1.64)	(1.78)	(2.18)	(0.71)	(0.64)
TLS	−0.004	0.005	0.004	0.046	0.037*	0.019
	(−0.30)	(0.43)	(0.26)	(1.36)	(1.97)	(0.63)
AQ × *TLS*	0.247**	−0.069	−0.083	1.151**	0.145	0.001
	(2.08)	(−0.58)	(−0.87)	(2.21)	(0.50)	(0.00)
BETA	0.013*	0.017***	0.007	0.01	0.014***	0.004
	(2.01)	(2.96)	(0.93)	(1.58)	(2.82)	(0.61)
ln(*SIZE*)	−0.014***	−0.003**	0.001	−0.013***	−0.002**	0.002
	(−5.20)	(−2.68)	(0.88)	(−4.92)	(−2.46)	(1.06)
ln(*BM*)	0.010**	0.015***	0.013***	0.008**	0.014***	0.013***
	(2.78)	(4.72)	(4.49)	(2.13)	(4.66)	(4.28)
MOM	−0.008	−0.007	0.002	0.003	0.000	0.005
	(−0.69)	(−0.55)	(0.15)	(0.41)	(−0.03)	(0.65)
YIELD	0.048	−0.021	−0.061	0.014	0.004	−0.045
	(0.56)	(−0.34)	(−0.77)	(0.18)	(0.06)	(−0.56)
IVOL	0.615***	0.399**	0.179	0.272	0.238	0.212
	(3.24)	(2.30)	(0.69)	(1.58)	(1.50)	(0.65)
PRICE	−0.074**	−0.035***	−0.012	−0.056	−0.031***	−0.009
	(−2.14)	(−3.54)	(−1.12)	(−1.68)	(−3.11)	(−0.83)
ILLIQ	0.055*	0.071	0.605***	0.053	0.062	0.606***
	(1.78)	(0.82)	(3.09)	(1.67)	(0.70)	(3.05)
Mean Adj. R^2	0.0799	0.0924	0.0907	0.0911	0.0933	0.0944

---

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed except for *AQ* × *TLS* (which is one-tailed).

This table presents the time-series mean coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions using January observations only. The regression is run separately for each institutional ownership (*IO*) tercile. The dependent variable is monthly returns. All explanatory variables are winsorized at the 1 percent and 99 percent levels; the fiscal year-end explanatory variables are mapped to returns three months later. The coefficients on *PRICE* and *ILLIQ* are multiplied by 100. t-statistics, shown in parentheses below the coefficient estimates, are based on Newey-West standard errors with two lags. The sample period runs from January 1971 to December 2006. All variables are defined in Appendix B.

quarters ending in December (regardless of whether their fiscal year ends in December—i.e., firms with upcoming annual or quarterly financial information in January) are dropped from the sample, as the *AQ* coefficient is similar in columns 1 and 3. Finally, the magnitude of the AQ premium is also unaffected if all observations with January earnings announcements, whether

TABLE 9

Testing the Information Hypothesis

$$January\ R_{it} = b_0 + b_1 AQ_{it-1} + b_2 Controls_{it-1} + \varepsilon_{it} \qquad (3)$$

	Full Sample (Same as Column 5 in Table 6) (1)	**Remove Obs. with Fiscal Years Ending in Dec.** (2)	**Remove Obs. with Fiscal Quarters Ending in Dec.** (3)	**Remove Obs. with Jan. Earnings Announcements** (4)
Intercept	0.059***	0.061***	0.064***	0.066***
	(8.42)	(7.47)	(6.13)	(9.73)
AQ	0.263***	0.281***	0.273***	0.271***
	(4.78)	(3.58)	(3.02)	(4.95)
BETA	0.011**	0.013***	0.012***	0.009**
	(2.50)	(2.92)	(2.89)	(2.11)
ln(SIZE)	−0.007***	−0.008***	−0.008***	−0.008***
	(−5.30)	(−5.10)	(−4.43)	(−4.84)
ln(BM)	0.016***	0.013***	0.011*	0.015***
	(5.28)	(3.41)	(1.79)	(4.10)
MOM	−0.015**	−0.016**	−0.026**	−0.02***
	(−2.36)	(−2.32)	(−2.58)	(−2.94)
YIELD	0.125***	0.133***	0.045	0.093**
	(3.01)	(2.85)	(1.07)	(2.36)
IVOL	0.699***	0.577***	0.658**	0.712***
	(5.19)	(3.10)	(2.16)	(4.39)
PRICE	−0.016	−0.032*	−0.044***	−0.036**
	(−1.25)	(−1.75)	(−2.76)	(−2.54)
ILLIQ	0.112***	0.101***	0.046	0.104***
	(3.68)	(3.54)	(1.33)	(3.04)
Mean Adj. R^2	0.1301	0.1242	0.1334	0.1325

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All significance tests are two-tailed.

This table presents the time-series mean coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions using January observations only. The dependent variable is monthly returns. All explanatory variables are winsorized at the 1 percent and 99 percent levels; the fiscal year-end explanatory variables are mapped to returns three months later. The coefficients on *PRICE* and *ILLIQ* are multiplied by 100. t-statistics, shown in parentheses below the coefficient estimates, are based on Newey-West standard errors with two lags. The sample period runs from January 1971 to December 2006. All variables are defined in Appendix B.

annual or quarterly, are dropped from the sample, as the *AQ* coefficient is similar in columns 1 and 4.

Alternative Measures of Accruals Quality

To shed light on the generalizability of our results for AQ, we examine two other commonly used measures of accruals quality: (1) the absolute value of performance-adjusted abnormal accruals (Kothari et al. 2005), and (2) the absolute value of the residual from the modified Dechow and Dichev (2002) model. The results are qualitatively similar to those for AQ—i.e., there is no annual premium in both cases, both proxies predict higher abnormal returns only in January, and the January premium in both cases is increasing in tax loss selling (untabulated). Thus, our results do not appear to be specific to the modified Dechow and Dichev (2002) AQ measure, but call into question the risk status of accruals quality in general.[23]

Controlling for Cash Flow Shocks

Since cash flow shocks affect realized returns and AQ is empirically highly correlated with negative cash flow shocks, Ogneva (2008) argues that controlling for cash flow shocks is important when realized returns are used as a proxy for expected returns in asset-pricing tests on AQ. To examine whether cash flow shocks affect our results, we follow Ogneva (2008) and control for such shocks using the scaled AQ measure (*SAQ*). *SAQ* is computed as AQ scaled by average absolute accruals estimated over years *t*–4 through *t* (the period over which AQ is measured).[24] In untabulated results, we find that the main results in Table 6 are insensitive to controlling for cash flow shocks; like AQ, *SAQ* predicts higher returns only in January.

Daniel and Titman (1997) Test

Finally, we use the Daniel and Titman (1997) approach to shed more light on whether AQ reflects risk or mispricing. Specifically, we examine the extent to which average returns are generated by the AQ factor loading (estimated using four-factor Fama-French regressions including the *AQFactor*) rather than the firm characteristic AQ.[25] If the AQ factor proxies for risk, then one would expect variation in the AQ factor loading rather than AQ to capture the cross-section of returns. Using 25 (5 × 5) portfolios sorted on AQ and the AQ factor loading, we find that increasing the AQ factor loading has no effect on average returns within any AQ portfolio. However, increasing AQ predicts returns for three of the five AQ factor loading portfolios (untabulated). Thus, any return premium to AQ appears to be driven by the AQ characteristic and not by sensitivity to an aggregate risk factor proxied by the AQ factor. Like our seasonality results, these findings are difficult to reconcile with AQ being a unique and priced risk factor.

VI. CONCLUSION

As an alternative way to shed light on the continuing debate over whether accruals quality is a priced risk factor, we examine the calendar-year pricing of the modified Dechow and Dichev (2002) accruals quality measure (AQ), a widely used proxy for accruals quality. Our results indicate that AQ is positively related to future abnormal stock returns only in January. Conversely, in non-January months, AQ is negatively correlated with future returns such that there is no AQ premium (typically) on an annual basis. Thus, any AQ premium appears entirely due to the month of January. Moreover, about half of the January premium occurs during the first five trading days in January. The January AQ premium is stable across time and across different sample periods, occurs for both small and large firms, and cannot be explained by the other January effects documented by prior literature. We also find that the January AQ premium varies positively with the likelihood of tax loss selling around the turn of the year, consistent with the premium reflecting, at least partly, the stock price effects of tax loss selling.

---

[23] The result for performance-adjusted abnormal accruals is not inconsistent with the well-known accrual anomaly; they predict negative returns on an annual basis and in the non-January period, but positive returns in January.

[24] Ogneva (2008) uses direct proxies as well as the scaled AQ measure to control for cash flow shocks. To be conservative, we use the scaled AQ measure because she seems to find stronger results with this measure (in support of AQ as a risk factor).

[25] Because we find no annual premium to AQ for our full sample period, we perform this test for the January 1985–November 2003 sample period, during which AQ does predict future returns on an annual basis. Note that unless an annual return premium is observed, the Daniel and Titman (1997) tests are pointless.

Taken together, these findings strongly suggest that the AQ premium is an artifact of mispricing around the turn of the year, rather than reflecting systematic information risk. Given traditional conceptions of risk in the existing literature, it is difficult to understand why such risk would be confined only to January, why half of it would manifest itself during the first trading week in January, why a January risk premium would then reverse itself during the rest of the year, and why the risk premium would be increasing in tax loss selling. Thus, although our findings are by no means conclusive evidence against a risk interpretation of AQ, any risk explanation should explain the nature of such risk and how it squares with the evidence in this study. We also caution that our results only cast doubt on whether accruals quality reflects systematic information risk; they do not in any way imply that other aspects of information risk, or other proxies for such risk, are not priced.

REFERENCES

Aboody, D., J. Hughes, and J. Liu. 2005. Earnings quality, insider trading, and cost of capital. *Journal of Accounting Research* 43 (5): 651–673.

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5: 31–56.

Armstrong, C., S. Banerjee, and C. Corona. 2009. Information quality and the cross-section of expected returns. Working paper, University of Pennsylvania, MIT, and University of Texas.

Ashbaugh-Skaife, H., D. Collins, W. Kinney, and R. LaFond. 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity. *Journal of Accounting Research* 47 (1): 1–43.

Athanassakos, G. 1992. Portfolio rebalancing and the January effect in Canada. *Financial Analysts Journal* 48 (6): 67–78.

Ball, R., S. P. Kothari, and J. Shanken. 1995. Problems in measuring portfolio performance: An application to contrarian investment strategies. *Journal of Financial Economics* 38: 79–107.

Barton, J., and M. Mercer. 2005. To blame or not to blame: Analysts' reactions to external explanations for poor financial performance. *Journal of Accounting and Economics* 39 (3): 509–533.

Berges, A., J. McConnell, and G. Schlarbaum. 1984. The turn-of-the-year in Canada. *Journal of Finance* 39 (1): 185–192.

Bernard, V., J. Thomas, and J. Wahlen. 1997. Accounting-based stock price anomalies: Separating market inefficiencies from risk. *Contemporary Accounting Research* 14 (2): 89–136.

Bharath, S., J. Sunder, and S. Sunder. 2008. Accounting quality and debt contracting. *Accounting Review* 83 (1): 1–28.

Bhardwaj, R., and L. Brooks. 1992. The January anomaly: Effects of low share price, transaction costs, and bid-ask bias. *Journal of Finance* 47 (2): 553–575.

Blume, M., and R. Stambaugh. 1983. Biases in computed returns: An application to the size effect. *Journal of Financial Economics* 12 (3): 387–404.

Botosan, C. 1997. Disclosure level and the cost of equity capital. *Accounting Review* 72 (3): 323–349.

Botosan, C., and M. Plumlee. 2002. A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research* 40 (1): 21–40.

Brauer, G., and E. Chang. 1990. Return seasonality in stocks and their underlying assets: Tax-loss selling versus information explanations. *Review of Financial Studies* 3 (2): 255–280.

Brickley, J., S. Manaster, and J. Schallheim. 1991. The tax-timing option and the discounts on closed-end investment companies. *Journal of Business* 64 (3): 287–312.

Brown, S., and S. Hillegeist. 2007. How disclosure quality affects the level of information asymmetry. *Review of Accounting Studies* 12 (2–3): 443–477.

Chang, E., and J. M. Pinegar. 1986. Return seasonality and tax-loss selling in the market for long-term government and corporate bonds. *Journal of Financial Economics* 17 (2): 391–415.

Chen, S., T. Shevlin, and Y. Tong. 2007. Does the pricing of financial reporting quality change around dividend changes? *Journal of Accounting Research* 45 (1): 1–40.

Christensen, P., L. de la Rosa, and G. Feltham. 2010. Information and the cost of capital: An ex-ante perspective. *Accounting Review* 85 (3): 817–848.

Constantinides, G. 1984. Optimal stock trading with personal taxes: Implications for prices and the abnormal January returns. *Journal of Financial Economics* 13 (1): 65–89.

Core, J., W. Guay, and R. Verdi. 2008. Is accruals quality a priced risk factor? *Journal of Accounting and Economics* 46 (1): 2–22.

Daniel, K., and S. Titman. 1997. Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance* 52 (1): 1–33.

Davis, J. 1994. The cross-section of realized stock returns: The pre-COMPUSTAT evidence. *Journal of Finance* 49 (5): 1579–1593.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *Accounting Review* 77 (Supplement): 35–59.

Doran, J., D. Jiang, and D. Peterson. 2010. Short-sale constraints and the idiosyncratic volatility puzzle: An event study approach. Working paper, Florida State University.

Doyle, J., W. Ge, and S. McVay. 2007. Accruals quality and internal control over financial reporting. *Accounting Review* 82 (5): 1141–1170.

Easley, D., S. Hvidkjaer, and M. O'Hara. 2002. Is information risk a determinant of asset prices? *Journal of Finance* 57 (5): 2185–2221.

Easley, D., and M. O'Hara. 2003. Asset pricing and market microstructure. In *Handbook of the Economics of Finance*, edited by G. Constantinides, M. Harris, and R. Stulz, 1021–1047. Oxford, U.K.: Elsevier Scientific Press.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* 59 (4): 1553–1583.

Ecker, F., J. Francis, I. Kim, P. Olsson, and K. Schipper. 2006. A returns-based representation of earnings quality. *Accounting Review* 81 (4): 749–780.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153–193.

Fama, E., and J. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–636.

Francis, J., A. Huang, S. Rajgopal, and A. Zang. 2008a. CEO reputation and earnings quality. *Contemporary Accounting Research* 25 (1): 109–147.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *Accounting Review* 79 (4): 967–1010.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39 (2): 295–327.

Francis, J., D. Nanda, and P. Olsson. 2008b. Voluntary disclosure, earnings quality, and cost of capital. *Journal of Accounting Research* 46 (1): 53–99.

Griffiths, M., and R. White. 1993. Tax-induced trading and the turn-of-the-year anomaly: An intraday study. *Journal of Finance* 48 (2): 575–598.

Grinblatt, M., and T. Moskowitz. 2004. Predicting stock price movements from past returns: The role of consistency and tax-loss selling. *Journal of Financial Economics* 71 (3): 541–579.

Gultekin, M., and N. Gultekin. 1983. Stock market seasonality: International evidence. *Journal of Financial Economics* 12 (4): 469–481.

Gultekin, M., and N. Gultekin. 1987. Stock return anomalies and the tests of the APT. *Journal of Finance* 42 (5): 1213–1224.

Huddart, S., and B. Ke. 2007. Information asymmetry and cross-sectional variation in insider trading. *Contemporary Accounting Research* 24 (1): 195–232.

Hughes, J., J. Liu, and J. Liu. 2007. Information asymmetry, diversification, and cost of capital. *Accounting Review* 82 (3): 705–729.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48: 65–91.

Kato, K., and J. Schallheim. 1985. Seasonal and size anomalies in the Japanese stock market. *Journal of Financial and Quantitative Analysis* 20 (2): 243–260.

Keim, D. 1983. Size-related anomalies and stock market seasonality: Further empirical evidence. *Journal of Financial Economics* 12 (1): 13–32.

Keim, D. 1986. Dividend yields and the January effect. *Journal of Portfolio Management* Winter 54–60.

Keim, D. 1989. Trading patterns, bid-ask spreads, and estimated stock returns: The case of common stocks at calendar turning points. *Journal of Financial Economics* 25: 75–98.

Kim, D. 2006. On the information uncertainty risk and the January effect. *Journal of Business* 79 (4): 2127–2162.

Kim, D., and Y. Qi. 2010. Accruals quality, stock returns, and macroeconomic conditions. *Accounting Review* 85 (3): 937–978.

Kothari, S. P., A. Leone, and C. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.

Kramer, C. 1994. Macroeconomic seasonality and the January effect. *Journal of Finance* 49 (5): 1883–1891.

Kravet, T., and T. Shevlin. 2009. Accounting restatements and information risk. *Review of Accounting Studies* 15 (2): 264–294.

Krishnan, G., B. Srinidhi, and L. Su. 2008. Inventory policy, accruals quality and information risk. *Review of Accounting Studies* 13 (2–3): 369–410.

Lakonishok, J., A. Shleifer, and R. Vishny. 1994. Contrarian investment, extrapolation, and risk. *Journal of Finance* 49 (5): 1541–1578.

Lambert, R., C. Leuz, and R. Verrecchia. 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45 (2): 385–420.

Lee, G., and R. Masulis. 2009. Seasoned equity offerings: Quality of accounting information and expected flotation costs. *Journal of Financial Economics* 92 (3): 443–469.

Loughran, T. 1997. Book-to-market across firm size, exchange, and seasonality: Is there an effect? *Journal of Financial and Quantitative Analysis* 32 (3): 249–268.

Lui, D., S. Markov, and A. Tamayo. 2007. What makes a stock risky? Evidence from sell-side analysts' risk ratings. *Journal of Accounting Research* 45 (3): 629–665.

Merton, R. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42 (3): 483–510.

Ng, L., and Q. Wang. 2004. Institutional trading and the turn-of-the-year effect. *Journal of Financial Economics* 74 (2): 343–366.

Ogden, J. 2003. The calendar structure of risk and expected returns on stocks and bonds. *Journal of Financial Economics* 70 (1): 29–67.

Ogneva, M. 2008. Accrual quality and expected returns: The importance of controlling for cash flow shocks. Working paper, Stanford University.

Poterba, J., and S. Weisbenner. 2001. Capital gains tax rules, tax-loss trading, and turn-of-the-year returns. *Journal of Finance* 56 (1): 353–368.

Reinganum, M. 1983. The anomalous stock market behavior of small firms in January: Empirical tests for tax loss selling effects. *Journal of Financial Economics* 12 (1): 89–104.

Reinganum, M., and P. Gangopadhyay. 1991. On information release and the January effect: Accounting-information hypothesis. *Review of Quantitative Finance and Accounting* 1 (2): 169–176.

Ritter, J. 1988. The buying and selling behavior of individual investors at the turn of the year. *Journal of Finance* 43 (3): 701–717.

Ritter, J., and N. Chopra. 1989. Portfolio rebalancing and the turn-of-the-year effect. *Journal of Finance* 44 (1): 149–166.

Roll, R. 1983a. On computing mean returns and the small firm premium. *Journal of Financial Economics* 12 (3): 371–386.

Roll, R. 1983b. Vas ist das? The turn-of-the-year effect and the return premia of small firms. *Journal of Portfolio Management* 9: 18–28.

Rozeff, M., and W. Kinney. 1976. Capital market seasonality: The case of stock returns. *Journal of Financial Economics* 3 (4): 379–402.

Sengupta, P. 1998. Corporate disclosure quality and the cost of debt. *The Accounting Review* 73 (4): 459–474.
Sias, R., and L. Starks. 1997. Institutions and individuals at the turn-of-the-year. *Journal of Finance* 52 (4): 1543–1562.
Srinidhi, B., and F. Gul. 2007. The differential effects of auditors' nonaudit and audit fees on accrual quality. *Contemporary Accounting Research* 24 (2): 595–629.
Tinic, S., and R. West. 1984. Risk and return: January vs. the rest of the year. *Journal of Financial Economics* 13 (4): 561–574.

APPENDIX A

A Selection of Published Studies Viewing AQ as a Proxy for Information Risk

The following selection of published articles in top-tier accounting journals shows that many researchers continue to view the modified Dechow and Dichev (2002) AQ measure as a proxy for information risk, despite the uncertainty about AQ's risk status. Articles are listed alphabetically.

Articles that (1) test and conclude that AQ proxies for information risk; (2) use AQ as a proxy for information risk in different settings; or (3) make inferences based on the assumption that AQ proxies for information risk are shown below.

- Aboody et al. (2005) examine whether earnings quality is priced (using AQ as one measure of earnings quality) and whether insiders earn greater profits when they trade on stocks with a higher exposure to the earnings quality risk factor. They conclude that they find "evidence consistent with pricing of the earnings quality factor and insiders trading more profitably in firms with higher exposure to that factor" (Aboody et al. 2005, 651).
- Ashbaugh-Skaife et al. (2009) examine whether ineffective internal control results in less reliable financial reporting, thus increasing the information risk faced by investors that manifests in a higher cost of equity. They state: "Prior research suggests that firms with low-quality accruals have a higher cost of equity (Francis et al. 2005)" (Ashbaugh-Skaife et al. 2009, 36). They examine whether poor internal controls have an incremental effect on the cost of equity after controlling for AQ.
- Bharath et al. (2008) analyze the impact of borrower accounting quality on debt contracting. Their composite measure of accounting quality incorporates AQ. They state that "an interpretation of our results could be that stringent contract terms for low accounting quality borrowers reflect lenders' compensation for information risk. This interpretation is consistent with Easley et al. (2002), Easley and O'Hara (2003), and Francis et al. (2005)" (Bharath et al. 2008, 4). They also mention that their evidence "further supports the hypothesis that information risk is a priced source of risk distinct from default risk" (Bharath et al. 2008, 5).
- Chen et al. (2007) examine whether accruals quality (proxied by AQ) is a priced information risk factor in a dividend change setting. They conclude: "Overall, we interpret our results as being consistent with investors treating the information risk associated with the precision of financial statement information as a priced risk factor, with both the precision and pricing changing in predictable directions around dividend changes" (Chen et al. 2007, 1–2).
- Ecker et al. (2006) analyze a returns-based measure of earnings quality estimated from firm-specific asset-pricing regressions augmented by an earnings-quality-mimicking factor (the AQ factor). They state: "Empirical support for the view that earnings quality as measured by accruals quality is priced is provided by Francis et al. (2005) . . . Thus, conditioning on theory that shows information uncertainty is a priced risk factor, and on Francis et al.'s (2005) empirical evidence that the market prices information risk as captured by earnings quality measured by AQ, we view e-loadings as capturing the sensitivity of stock returns to earnings quality" (Ecker et al. 2006, 750–751).

- Francis et al. (2008a) examine the link between CEO reputation and earnings quality and find that firms with reputed CEOs have poorer earnings quality. They note the following implication of this finding: "Francis et al. (2005) and Aboody et al. (2005) argue that poor earnings quality, as proxied by the Dechow and Dichev (2002) measure of earnings quality [i.e., AQ] and absolute discretionary accruals, is associated with a higher cost of equity and debt capital and greater rents appropriated by insiders. On the basis of these findings, we argue that if reputed CEOs influence earnings quality to make reported earnings look good to maintain their reputations, cost of capital and insider rent appropriation increase" (Francis et al. 2008, 142).
- Francis et al. (2004) examine the relation between the cost of equity capital and seven earnings attributes, including accruals quality (proxied by AQ). They conclude: "On the whole, the weight of the evidence suggests that, among the seven attributes we consider, accrual quality is the dominant attribute in terms of cost of equity effects" (Francis et al. 2004, 1006).
- Francis et al. (2005) examine whether investors price accruals quality, a proxy for information risk. They state: "Measuring accruals quality (AQ) as the standard deviation of residuals from regressions relating current accruals to cash flows, we find that poorer AQ is associated with larger costs of debt and equity" (Francis et al. 2005, 295).
- Francis et al. (2008b) examine the relation between voluntary disclosure and earnings quality, and the pricing effects of voluntary disclosure unconditionally and conditional on earnings quality. Their composite measure of earnings quality incorporates AQ. They conclude: "On the whole, our results show that earnings quality has a first-order effect on cost of capital, while voluntary disclosure has no significant and distinct influence on the cost of capital" (Francis et al. 2008, 81).
- Kim and Qi (2010) examine whether AQ is a priced risk factor and how the pricing of AQ varies with macroeconomic conditions. They conclude that their results "suggest that AQ contributes to the cost of equity capital and that its pricing effect is associated with fundamental risk" (Kim and Qi 2010, 937).
- Kravet and Shevlin (2009) examine the association between accounting restatements and the pricing of information risk. They use AQ as a proxy for information risk and the discretionary component of AQ as a proxy for discretionary information risk. They "find that the negative market reaction resulting from restatements documented in prior literature is associated with the increase in the pricing of discretionary information risk" (Kravet and Shevlin 2009, 3).
- Krishnan et al. (2008) investigate whether inventory policy affects information risk that is priced by the market after controlling for AQ. They assert that "Francis et al. (2004, 2005) show that accruals quality is priced by the market as a systematic factor and that the market demands higher returns from firms with poorer accruals quality, consistent with Easley and O'Hara (2004)" (Krishnan et al. 2008, 371). The authors conclude that their "results imply that investors price the inventory policy as a proxy for unobservable underlying systematic risk that is incremental to information risk captured by accruals quality [AQ]" (Krishnan et al. 2008, 397).
- Lee and Masulis (2009) investigate the impact of accruals quality (proxied by AQ) on underwriter gross spreads during SEOs. They argue that "poor accruals quality by an issuer leads to high information risk, because it raises investor uncertainty and asymmetric information between the issuer and outside investors, which SEO announcements could exacerbate" (Lee and Masulis 2009, 452).
- Lui et al. (2007) examine whether certain firm characteristics, including earnings quality, are determinants of analysts' risk ratings. In motivating the use of earnings quality (proxied by

> AQ) as a determinant of the risk ratings, they note that "firms with lower earnings quality have been shown to have higher cost of equity (e.g., Francis et al. [2004, 2005])" (Lui et al. 2007, 654). The authors conclude that "idiosyncratic risk, size, book-to-market, leverage, accounting losses, and earnings quality are important determinants of analysts' risk ratings. We conclude that financial analysts gather and process information about investment risk in determining the risk ratings" (Lui et al. 2007, 659).

Studies that do not rely on AQ as a proxy for information risk but cite Francis et al.'s (2004, 2005) conclusion that information risk is priced are shown below.

- Barton and Mercer (2005): "Recent evidence suggests that information risk is not fully diversifiable and thus affects the firm's cost of capital (Botosan 1997; Easley et al. 2002; Francis et al. 2004, 2005)" (Barton and Mercer 2005, 511).
- Brown and Hillegeist (2007): "A growing body of literature reports a negative relation between various measures of disclosure quality and cost of capital estimates (Botosan 1997; Botosan and Plumlee 2002; Francis et al. 2005; Sengupta 1998)" (Brown and Hillegeist 2007, 445–446).
- Huddart and Ke (2007): "Another source of information risk may be related to earnings quality. Francis et al. (2005) present evidence that accruals quality is a priced risk factor" (Huddart and Ke 2007, 229).
- Srinidhi and Gul (2007): "Francis et al. (2005) show that firms with low accrual quality face higher costs of capital because of the need to compensate for the increased information risk that investors face" (Srinidhi and Gul 2007, 625).

APPENDIX B

Variable Measurement

Variable **Description**

AQ = modified Dechow and Dichev (2002) AQ measure, calculated by estimating a regression of total current accruals on lagged, current, and future cash flows, plus the change in revenue and the level of property, plant, and equipment. The regression is estimated in the cross-section each year for each industry with at least 20 observations. *AQ* is the standard deviation of firm-specific residuals over a five-year period, ending at the previous fiscal year-end;

SIZE = market value of common equity (in millions), calculated on the last trading day of the previous month;

BETA = calculated using a market-model regression (with value-weighted market returns) estimated over the 60 months prior to the current month;

BM = book-to-market ratio on the last trading day of the previous month, using book value at the previous fiscal year-end and market capitalization on the last trading day of the previous month;

MOM = calculated as the cumulative raw return from month $t-12$ to $t-2$ where t is the portfolio formation month;

YIELD = annual dividends per common share divided by the fiscal year-end stock price, calculated at the previous fiscal year-end;

IVOL = calculated as the standard deviation of the residuals from a market-model regression (with value-weighted market returns) estimated over the 36 months prior to the current month;

PRICE = stock price on the last trading day of the previous month;

ILLIQ = annual average of the ratio of the daily absolute return to the daily dollar trading volume, measured at the end of the previous calendar year;

TLS = tax loss selling proxy: either *ARET* or *PTLS* (defined below);

ARET = firm's buy-and-hold return over the previous calendar year, excluding the last five trading days in December, the measure is multiplied by -1;

PTLS = ratio of the stock price on the last trading day of the previous calendar year (excluding the last five trading days) and the maximum stock price between January 1 and the last trading day of the year (excluding the last five trading days), the measure is multiplied by -1; and

IO = sum of all institutional holdings of the stock scaled by the number of shares outstanding. We use the most recent institutional ownership data for a firm prior to December 1 of the previous year. The number of shares outstanding is measured at the end of the month in which the ownership data are reported.

CALCUTTA UNIVERSITY CENTRAL LIBRARY

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1383–1414

American Accounting Association
DOI: 10.2308/accr-10034

What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions

Dawn Matsumoto
University of Washington

Maarten Pronk
Erik Roelofsen
Erasmus University Rotterdam

ABSTRACT: Conference calls held in conjunction with an earnings release have become increasingly common in recent years, yet there is little evidence regarding the reasons that these calls are incrementally informative over the accompanying press release. Using a sample of more than 10,000 conference-call transcripts, we examine the information content of both segments of the call—the presentation and the discussion segment. We find that both segments have incremental information content over the accompanying press release. However, discussion periods are relatively more informative than presentation periods, and this greater information content is positively associated with analyst following. We also find that managers provide increased disclosures during the presentation segment when firm performance is poor, but relatively more information is released during discussion periods in these circumstances. Overall, our results are consistent with the notion that active analyst involvement in conference calls increases the information content of the calls, particularly when firm performance is poor.

Keywords: *conference calls; voluntary disclosure; financial analysts; information content; information environment.*

We thank Beverly Walther (associate editor), and two anonymous reviewers for their comments and suggestions. We also thank Willemijn van Eijck and Rosiandra de Jesus for their valuable research assistance. The comments of Orie Barron, Bob Bowen, Dave Burgstahler, Andy Call, Shuping Chen, Weili Ge, Bin Ke, Bill Mayew, Jim McKeown, Shiva Rajgopal, Terry Shevlin, Mark Soliman, Lloyd Tanlu, Jim Vincent, and workshop participants at the Free University Amsterdam, INSEAD, The Pennsylvania State University, Tilburg University, The University of Texas, and University of Washington are greatly appreciated. Professor Matsumoto acknowledges the financial support of the University of Washington the William R. Gregory Accounting Fellowship, the Emmett S. Harrington Endowed Professorship in Accounting, and the Accounting Development Fund. Professor Pronk acknowledges financial support of the Netherlands Organization for Scientific Research (NWO R46–579).

Editor's note: Accepted by Beverly Walther.

Submitted: July 2008
Accepted: January 2011
Published Online: April 2011

I. INTRODUCTION

Over the past decade, conference calls have become an increasingly common form of voluntary disclosure (Bushee et al. 2003). Prior studies suggest that these calls are significant information events to the market (Frankel et al. 1999; Bowen et al. 2002; Kimbrough 2005) and recent proposed changes by the Public Company Accounting Oversight Board (PCAOB) suggest that these calls can also be informative for assessing auditing risk (PCAOB Release 2009-007). However, the reasons conference calls are informative have not been widely investigated. There are at least two reasons that conference calls might be incrementally informative over a press release. First, managers are able to provide information in a less constrained fashion relative to financial statements and written press releases. Second, analysts can play a direct role in uncovering information during the question-and-answer session (hereafter, the "discussion" portion of the call), asking follow-up questions, requesting more detail, and perhaps questioning management's interpretation of events. The purpose of this study is to investigate (1) whether each of these two main components of earnings-related conference calls are incrementally informative to the market and, if both are informative, which segment has greater information content; and (2) whether the information content of each segment is related to firm performance.

During the presentation portion of the call, managers provide their interpretation of the firm's performance during the quarter and provide any additional, voluntary disclosures they wish to communicate. In addition to possibly providing new disclosures during the presentation, managers also provide the information verbally, which is potentially informative to the market because of the information content of verbal cues (Mayew and Venkatachalam 2009). If the ability to disclose information in a less constrained fashion results in greater disclosure, then we would expect the presentation portion of the call to be incrementally informative over the accompanying press release. We would also expect the discussion portion of the call to be incrementally informative if analysts' ability to question management's interpretation of events and/or to elicit additional information from managers is informative to the market. Whether the primary advantage of conference calls comes from expanded disclosures during the presentation or from managers' interaction with analysts is an empirical question.

Our analyses are based on transcripts of more than 10,000 earnings-announcement-related conference calls held during regular trading hours for the period January 2003 to December 2005. Using these transcripts, we compute approximate start and end times for each portion of the call and then use intra-day trading data to calculate absolute returns during each segment. We use this measure as our proxy for information content. However, because absolute returns during the relatively narrow windows associated with each segment of the call are potentially noisy measures of information content, we supplement some of our tests with analyses based on call length (in number of words) as an alternative measure of information content.

We first examine the incremental information content of each segment of the call and find that both the presentation and discussion have incremental information content over the accompanying press release. However, we find statistically greater abnormal absolute returns during the discussion portion of the call relative to the presentation, suggesting that the discussion portion is relatively more informative. Consistent with the explanation that greater analyst involvement is the source of the relatively greater information content of the discussion portion of the call, we find that the informativeness of the discussion session increases with analyst following, while the informativeness of the presentation does not. As a result, the *relative* information content of the discussion session increases in analyst following. To further support our conjecture, we also examine the *length* of the discussion session relative to the presentation and find that the relative length also increases with analyst following.

Although these results support the notion that the primary benefit of conference calls comes from analyst involvement during the discussion session, we recognize that our analyses are tests of association and not tests of causation. Thus, it is inappropriate to characterize our results as providing evidence that analysts are the *cause* of the relatively greater information content during the discussion portion of the call, although the results are consistent with such an interpretation.

We next examine whether the incremental information content of each segment of the call is related to firm performance. On the one hand, disclosure theories generally predict a bias toward the disclosure of good news, with many prior empirical studies finding a positive relation between firm performance and disclosure (Lev and Penman 1990; Lang and Lundholm 1993; Miller 2002; Frankel et al. 1999). However, other studies find that managers are more likely to preempt bad news (Skinner 1994; Soffer et al. 2000), and that they provide greater supplemental disclosures with bad news forecasts (Baginski et al. 2004). Thus, prior research is mixed as to whether managers provide more or less disclosure when performance is poor. Assuming that the information content of the presentation portion of the call is more representative of managers' initial disclosure incentives while the information content of the discussion portion represents disclosures made as a result of analyst questioning, we add to the understanding of managers' disclosure incentives by examining each segment of the conference call separately.

Our results suggest that managers provide greater disclosures during both the presentation and discussion segments when performance is poor. However, we find that the information content of the discussion portion of the call *relative* to the presentation portion of the call is greater when performance is poor, suggesting that although managers provide greater disclosures when the firm has not performed well, the disclosures made at the behest of analysts are more informative in these circumstances. We find similar relations using length in words as an alternative measure of information content.

We find these relations after controlling for potential economic events that might lead to more informative presentations, such as the magnitude of special items, merger and restructuring activities, and the absolute returns over the quarter. We also run our regressions including firm fixed effects, thereby controlling for firm-specific factors that might potentially affect the informativeness of the presentation and discussion portions of the conference call.

Finally, to further substantiate that our evidence on the relation between the information content of presentation and discussion sessions and firm performance are due to real changes in the content of disclosures (as opposed to different market reactions to similar disclosures), we examine the relation between firm performance and the *type* of disclosures made during presentations and discussions. Focusing on two particular dimensions of content, financial (versus nonfinancial) and forward-looking (versus backward-looking), we find that managers provide less financial and more future-oriented disclosures in the presentation when firm performance is poor. Since the typical earnings release focuses on financial, backward-looking information, it is likely that nonfinancial, future-oriented disclosures are incrementally informative to the market. Thus, this result is consistent with our finding that presentations are more informative when firm performance is poor.

Our study contributes to the literature in several ways. It is one of the first studies to separately examine the information content of the two segments of conference calls, adding to the understanding of the benefits of conference calls as a disclosure mechanism. Our evidence suggests the benefit of conference calls is not solely due to managers' ability to verbally provide expanded disclosures in a less structured environment, but is also due to analysts' active involvement in the call, implying that analysts play a role in shaping *public* disclosures made by the firm. These results are also consistent with recent research by Mayew and Venkatachalam (2009) who find that verbal cues during the discussion portion of the call are informative to the market, particularly when analyst scrutiny is high.

Second, our study is one of the first to examine actual conference-call transcripts to examine the amount and type of disclosures made during conference calls and, more specifically, in the two segments of the call.[1] Prior studies have generally only examined the *existence* of conference calls as a measure of firm disclosure (Frankel et al. 1999; Tasker 1998; Bowen et al. 2002) because data on the actual content of calls were not readily available. By using quarter-specific call transcripts, we are able to examine variations in disclosures made by a given firm across time, providing new insights into managers' disclosure decisions. For example, the determinants of the decision to begin hosting conference calls are not necessarily the same as the determinants of the quantity and type of disclosures to provide during a call once the decision to host a call has been made. Thus, while prior studies find that firms that host conference calls are more profitable than those that do not (Frankel et al. 1999), we find that, conditional on hosting a call, firms disclose more information during calls when earnings performance is *poor*. Moreover, while our result is consistent with other research suggesting managers voluntarily provide expanded disclosures in the presence of bad news (Baginski et al. 2004), the fact that we find even greater increases in disclosures during the *discussion* (relative to the presentation) segment of the call suggests that some of the increased disclosures are not completely voluntary. In other words, some disclosures would perhaps not have been made were it not for questioning by analysts.

Our findings are also relevant to recent research by Roychowdhury and Sletten (2011), who find that more of the overall news released during a quarter is concentrated around the earnings announcement in quarters when the overall news is negative. They argue that this concentration is due to the fact that managers have incentives to delay the disclosure of bad news—that is, rather than make preemptive disclosures, managers wait until the earnings announcement to disclose bad news. Our findings suggest that another reason for the greater concentration of news around the earnings announcement may be due to increased disclosures made by managers as a result of analyst inquiries during the discussion portion of the conference call.

In the next section we discuss the prior literature and our hypotheses. We discuss our sample and variable measurement in Section III. Section IV presents our analysis of the information content of the presentation and discussion segments of the call. Section V presents additional analyses on the type of disclosures made during presentations and discussion periods. Section VI concludes.

II. PRIOR LITERATURE AND HYPOTHESIS DEVELOPMENT

Information Content of Presentations and Discussion Periods

A number of prior studies find evidence suggesting that conference calls held in conjunction with an earnings announcement are incrementally informative over the accompanying press release. Frankel et al. (1999) document abnormal trading volume and return volatility during the call period. Bowen et al. (2002) find that analyst forecast accuracy increases after earnings announcements that include conference calls relative to those that do not. Kimbrough (2005) finds that the market under-reacts less to current earnings when conference calls are held in conjunction with an earnings announcement relative to when a call is not hosted.

The reasons that conference calls are incrementally informative could reflect several factors. First, it is possible that managers voluntarily provide new information in the conference call relative to the press release. Frankel et al. (1999, 36) assert that "conference calls, being less formal than written press releases, are subject to a lower standard of legal liability than statements made during press releases." If managers believe this to be true, then they are likely to be willing to provide certain information during the presentation portion of the conference call that they would not be willing to provide in a press release, making the call incrementally more informative.

---

[1] In concurrent work, Frankel et al. (2009) examine the impact of missing analysts' expectations by a penny on the length and tone of conference calls and on the probability of issuing forward-looking guidance. Consistent with our results, they find that firms that miss analyst forecasts have longer conference calls. Our study differs from theirs in that we focus on the relation between information provided in the two segments of the call (presentation and discussion periods) whereas they focus primarily on the relation between missing analysts' expectations and overall call length and tone (both presentations and discussions combined).

Second, conference calls are spoken disclosures (versus written), such that it is possible that managers' verbal cues are incrementally informative to call participants. Mayew and Venkatachalam (2009) investigate verbal cues in conference calls and find evidence of a relation between the verbal effect displayed by managers on a call and market returns.

These two reasons suggest that the presentation portion of the conference call will have incremental information content over the accompanying press release. Verbal cues will also potentially impact the informativeness of the discussion portion of the call because of the spontaneous nature of questions and answers.[2] In addition, because of analysts' questions, managers could reveal information that they might not otherwise have disclosed. There are at least two reasons that managers might not initially disclose certain information. First, managers may be unaware that a particular piece of information is important. Second, managers may not initially reveal certain information because they would prefer not to. Disclosure models generally predict that managers will disclose good news and withhold bad news. However, the failure to disclose information will be inferred as bad news, which can lead to an unraveling of the information, unless there are reasons for non-disclosure such as proprietary costs (Verrecchia 1983) or uncertainty about the arrival of information (Dye 1985). If managers fail to disclose certain information in the press release or the presentation portion of the call, then the market may not necessarily be able to infer the lack of disclosure as bad news because it is possible that the manager did not know the market wanted this information. However, if analysts directly ask managers for this information in the discussion portion of the conference call, then this "excuse" for non-disclosure is not viable and the manager will either disclose the information or the market will infer negative information from the unwillingness to disclose (Hollander et al. 2010). In general, prior research is consistent with analysts benefitting from the ability to ask questions (Mayew 2008; Mayew et al. 2009; Libby et al. 2008). For these reasons, we expect the discussion portion of the call to be incrementally informative as well.

Thus, we expect both the presentation and the discussion portions of the call to be incrementally informative over the accompanying press release. In addition, while there are reasons supporting the incremental informativeness of the presentation portion of the call, we expect the discussion session to have greater information content than the presentation segment. Mayew and Venkatachalam (2009) argue that verbal cues are likely to be most prominent during the discussion portion of the call. In addition, prior evidence on the benefits to analysts of participating in conference calls (Mayew 2008; Mayew et al. 2009; Libby et al. 2008) suggests this portion of the call is likely to be more informative than the presentation.

To the extent that analysts' involvement in conference calls leads to more informative discussion sessions, we expect the information content of discussion sessions to increase with analyst following. In addition, because it is likely to be more costly to determine the appropriate disclosures to make when analyst following is greater, managers possibly limit disclosures during the presentation and instead wait to hear analysts' questions during the discussion. Both these factors would result in a positive relation between analyst following and the relative information content of the presentation versus the discussion session.

---

[2] In fact, the evidence in Mayew and Venkatachalam (2009) is based on an analysis of speech during the discussion portion of the call. However, it is still possible that managers' verbal cues also impact the information content of the presentation portion of the call.

Information Content and Firm Performance

Evidence from prior studies on the relation between firm performance and disclosure is somewhat mixed. Some studies find a positive relation between performance and disclosure (Lev and Penman 1990; Lang and Lundholm 1993; Frankel et al. 1999; Miller 2002).[3] However, other studies find managers are more likely to preempt bad news (Skinner 1994; Soffer et al. 2000). Studies on supplemental disclosures made with earnings forecasts also provide mixed evidence. Baginski et al. (2004) find that managers provide greater supplemental disclosures when forecasting bad news. In contrast, Hutton et al. (2003) find similar levels of qualitative "soft talk" disclosures for good and bad news forecasts, but find greater levels of verifiable forward-looking statements with good news forecasts. The latter finding is consistent with experimental evidence provided in Hirst et al. (2008) that suggests managers have incentives to provide more supplemental disclosures when forecasting *good* news in order to increase the credibility of the forecast.

It is possible that managers provide more disclosures during the presentation portion of the call when firm performance is poor because poor performance cannot persist indefinitely. As a result, market participants will likely demand information about the reasons for the past poor performance (e.g., whether it is transitory) and plans for addressing the poor performance going forward. In other words, uncertainty is higher when performance is poor and managers might seek to alleviate this uncertainty with greater disclosure (we label this the "uncertainty argument"). However, it is also possible that managers have greater incentives to provide expanded disclosures when firm performance is good in order to increase the credibility of the news (we label this the "credibility argument"). Thus, we do not predict a directional relation between firm performance and the information content of the presentation portion of the call.

The above discussion also has implications for the relation between the information content of the discussion portion of the call and firm performance. On the one hand, the uncertainty argument suggests greater disclosure during the discussion session when performance is poor. Because managers likely make more significant changes to their business when performance is poor, analysts will likely need more information from managers in order to appropriately revise their forecasting models. If managers do not know the exact information that analysts need to make these adjustments, then they will not provide this information in the presentation and analysts will likely demand additional disclosures during the discussion period. Alternatively, the credibility argument suggests analysts demand additional disclosures from managers when performance is good in order to assess the credibility of the firm's performance. This latter argument, however, assumes that managers do not *voluntarily* provide these additional disclosures in the presentation portion of the call. As it seems likely that managers would voluntarily provide such disclosures, we conjecture that the uncertainty argument will have greater bearing on the relation between the information content of the discussion and firm performance. Thus, we predict the information content of discussion periods to be greater when firms report poor performance.

The fact that managers may not know the appropriate information to disclose during the presentation when performance is poor also suggests that the information content of the discussion session relative to the presentation will be negatively related to firm performance. That is, when firm performance is poor, the discussion session will be relatively more informative than the presentation. It is also possible that managers know the appropriate disclosures but do not disclose them voluntarily. As discussed previously, disclosure theory suggests that managers have incentives to delay the disclosure of bad news (Verrecchia 1983; Dye 1985) but are less able to credibly avoid these disclosures when directly questioned by analysts. To the extent there are more likely to be additional disclosures that managers fail to disclose during the presentation portion of the call but that analysts uncover during the discussion session, we also expect the relative informativeness of discussion sessions to increase when performance is poor.[4]

---

[3] While prior studies suggest that, on average, firms that host conference calls have better performance, this result does not necessarily imply that a given firm is more likely to host calls in quarters with better performance. More likely, the result is driven by the fact that firms are more likely to adopt a *policy* of hosting calls when performance is better. A firm that has established a policy of hosting conference calls following its earnings announcements would likely find it very costly not to host a call in quarters when earnings performance is poor.

III. DATA AND VARIABLE MEASUREMENT

Sample Selection

We begin our sample selection process by identifying announcements of earnings-related conference calls (via Dow Jones Calendar of Corporate Events) held during trading hours (calls beginning after 9:30 a.m. and before 2:30 p.m. EST) by NYSE/NASDAQ firms between January 1, 2003 and December 31, 2005. This process yields 19,043 potential conference calls. We limit our sample to calls held during trading hours because we use intra-day trading data to measure the information content of each segment of the call.[5] We then gather transcripts of calls from Voxant FD wire available through Factiva. We are unable to find transcripts for 5,236 calls, leaving us with a sample of 13,807 conference-call transcripts.[6] We lose 2,581 observations due to missing Trade and Quote (TAQ) data, I/B/E/S analyst forecast data, Compustat data, and/or CRSP data. Finally, because certain of our analyses include firm fixed effects, we eliminate firms with less than four conference calls (1,164 observations). Our final sample comprises 10,062 firm-quarters. Table 1, Panel A details the sample selection process.

Table 1, Panel B presents the sample distribution across calendar quarters. The most observations are in 2004, followed by 2005, and then 2003. The fact that the sample declines in 2005 appears to be due to missing transcripts on Factiva on a few days.[7] Panel C shows the

---

[4] The impact of the credibility argument would also result in a negative relation between firm performance and the relative information content of the discussion. If managers seek to enhance the credibility of good firm performance by providing expanded disclosures during the presentation, then the relative information content of the presentation versus the discussion will increase with firm performance (i.e., the information content of the discussion session relative to the presentation will be lower when firm performance is good).

[5] Of the initial sample of 36,074 calls, 39.6 percent of firms always hold calls during trading hours, 35.4 percent always hold calls after trading hours, 14.6 percent switched once during the three-year period either from during to after or vice versa, and 10.4 percent switched more than once. Thus, it appears that the decision to hold calls during or after trading hours tends to be a policy decision that is firm-specific. We compared the four groups of firms on several dimensions including size, analyst following, market-to-book, leverage, performance, and length of presentation and discussion. The only differences that we find are: (1) firms that always hold calls during trading hours are larger and have more leverage than firms that always hold calls after trading hours; (2) firms that hold calls after trading hours have higher market-to-book ratios than firms that switch once. Because there does not appear to be significant differences between the groups, we do not believe that our choice to examine only calls held during trading hours biases our results. However, it may limit the generalizability of our results to the extent that firms that host calls after hours behave differently than firms that host calls during trading hours.

[6] We suspect that the availability of transcripts on Voxant may be related to firm size and analyst coverage. We compare firms with transcripts available on Voxant to those without and indeed find that firms with transcripts are larger (in terms of assets and market capitalization) and more highly followed. This fact does not bias our results but may limit the generalizability of our results to the extent that larger firms behave differently than smaller firms in the way they disclose information, similar to the impact of our choice to limit our sample to calls held during trading hours (see footnote 5).

[7] For 2005, transcripts appear to be missing on April 22, April 26, and July 28. Thus, April and July show a lower number of calls compared to 2004. If we exclude these two months, the number of observations in 2005 is greater than in 2004. The decline in 2005 is not the result of firms hosting fewer conference calls in 2005—the number of conference calls listed on Dow Jones calendar of corporate events shows a continuous increase over the three years (11,046 calls hosted in 2003, 11,880 in 2004 and 13,148 in 2005). We do not believe these missing transcripts introduce any systematic bias into our analysis.

TABLE 1

Sample Selection and Descriptive Data

Panel A: Sample Attrition

Earnings-related conference calls held during trading hours	19,043
Transcripts not available on Voxant FD Wire	(5,236)
Missing data on TAQ or stock price < $1.00	(477)
Missing analyst forecast data on I/B/E/S	(1,753)
Missing Compustat data	(305)
Missing CRSP data	(46)
Less than four conference calls per firm	(1,164)
Final Sample	10,062

Panel B: Sample Distribution across Calendar Quarters

Year		No. of Obs.	% of Obs.
2003	January–March	669	6.65
2003	April–June	618	6.14
2003	July–September	808	8.03
2003	October–December	773	7.68
2004	January–March	898	8.92
2004	April–June	1,013	10.07
2004	July–September	1,001	9.95
2004	October–December	830	8.25
2005	January–March	1,015	10.09
2005	April–June	821	8.16
2005	July–September	771	7.66
2005	October–December	845	8.40
		10,062	100.00

Panel C: Calls per Firm

Calls Per Firm	No. of Firms	Total Calls
4	159	636
5	125	625
6	135	810
7	196	1,372
8	205	1,640
9	198	1,782
10	168	1,680
11	103	1,133
12	32	384
	1,321	10,062

(continued on next page)

TABLE 1 (continued)

Panel D: Distribution of Observations across Industry

NAICS[a] Industry	No. of Obs.	% of Obs.	% of Population[b]
Manufacturing	3,552	35.30	34.11
Finance and Insurance	1,814	18.03	28.09
Information	645	6.41	10.37
Retail Trade	558	5.55	3.96
Mining	488	4.85	3.12
Professional, Scientific, and Technical Services	391	3.89	4.16
Utilities	336	3.34	2.26
Transportation and Warehousing	300	2.98	2.36
Wholesale Retail	295	2.93	2.58
Administrative and Support, Waste Management, and Remediation Services	288	2.86	1.75
Accommodation and Food Services	236	2.34	1.67
Health Care and Social Assistance	213	2.12	1.41
Construction	179	1.78	1.02
Real Estate and Leasing	143	1.42	1.51
Other Services	87	0.86	0.38
Art, Entertainment, and Recreation	73	0.72	0.42
Educational Services	69	0.69	0.31
Agriculture, Forestry, Fishing, and Hunting	13	0.13	0.31
Unclassified Establishments	9	0.09	0.21
Missing NAICS Code	373	3.71	—
	10,062	100.00	100.00

[a] NAICS stands for the North American Industry Classification System.
[b] "% of Population" is the percent of firms on the NYSE and NASDAQ in each NAICS industry classification.

distribution of calls per firm across our sample. We have all 12 conference calls (four per year for three years) for only a small fraction of firms. The reasons for this include missing Compustat, I/B/E/S, CRSP, or TAQ data; missing transcripts on Voxant; calls held outside of trading hours; and no call listed on the Dow Jones Calendar of Corporate Events for a particular quarter (due either to missing data or the initiation/discontinuation of conference calls).[8]

Finally, Table 1, Panel D presents the distribution of our sample across industries, using the North American Industry Classification System (NAICS). We also report the percent of the population of NYSE and NASDAQ firms comprising each industry classification. Our sample spans a wide range of industries and, in general, reflects the general distribution of industries in the population. The one exception is the Finance and Insurance industry, which comprises only 18 percent of our sample but comprises 28 percent of the population of NYSE/NASDAQ firms.

---

[8] If we had 12 calls for all 1,321 firms in our sample, we should have 15,852 observations. Of the 5,790 quarters missing from our sample, 4 percent are due to missing Compustat, I/B/E/S, CRSP, or TAQ data, 29 percent are due to missing transcripts, 15 percent are due to calls held outside trading hours (see footnote 5), and 52 percent are due to no calls reported on the Dow Jones Calendar of Corporate Events in that quarter. Of the latter group, it appears that 39 percent are due to the initiation or discontinuation of calls (i.e., calls are held either before or after the missing quarter date but not both); whereas 61 percent (1,825 quarters) have calls held both before and after the missing quarter date. We randomly selected 100 of these firms (193 firm-quarters) and searched earnings-announcement-related press releases to find reference to a conference call. In 94 percent of the cases, we were able to find reference to a conference call suggesting that the Dow Jones Calendar of Corporate Events has missing observations and *not* that these firms strategically decided against hosting a call. Thus, it appears that hosting conference calls tends to be a policy choice for the vast majority of firms.

Measures of Information Content

To measure our dependent variables, we split the transcripts into the presentation and discussion portion of the call by searching for the first occurrence of the word "operator" after the first 1,500 characters of the transcript. We then check whether the transcripts were split appropriately by searching the first 400 characters of the discussion for keywords or phrases that are typically associated with the start of the discussion—for example, "instructions," "question-and-answer," "the floor is now open," and "at this time." For discussion transcripts that did not include any of the keywords, we performed the split manually. In addition, we remove any legal disclaimers at the end of each transcript.

Our primary measure of information content is based on the abnormal absolute returns during each portion of the call. Absolute returns have been used in other studies using intra-day data as a measure of return volatility (Frankel et al. 1999; Lee et al. 1994), and return volatility is a valid measure of information content (Holthausen and Verrecchia 1990).[9] In order to calculate this variable, we first need to determine the start and end times of each call segment. To this end, we use a subsample of 1,263 conference calls made between January and March 2005 for which we have exact start and end times for each segment of the call. We use these times to calculate the number of words spoken per minute during the presentation and discussion segments and use the median (160 and 157 for the presentation and discussion, respectively) to compute the approximate length, in minutes, of the presentation ($\#MIN^{PRES}$) and discussion ($\#MIN^{DISC}$) for our full sample.[10] We then assume that (1) the start of the presentation occurs 116 seconds after the scheduled start time of the conference call, which is the average time spent on the introduction to the conference call based on our subsample, and (2) the beginning of the discussion starts 28 seconds after the end of the presentation, which is the average time spent on instructions by the operator in our subsample.[11] End times are then determined based on the $\#MIN^{PRES}$ and $\#MIN^{DISC}$.

Using these start and end times, we calculate the absolute returns during the presentation ($|RET^{PRES}|$) and the discussion ($|RET^{DISC}|$) using TAQ data. Specifically, $|RET^{PRES}|$ ($|RET^{DISC}|$) is the absolute value of the difference in quote midpoints at the start and end of the presentation (discussion), scaled by the quote midpoint at the start of the presentation (discussion). Then, to control for potential patterns in intra-day trading, we subtract the median absolute return observed during the same time period as the presentation and discussion but on all non-conference call days during the quarter ($|ABRET^{PRES}|$ and $|ABRET^{DISC}|$).[12] The difference in returns, *DIFFRET*, is defined as $|ABRET^{PRES}|$ less $|ABRET^{DISC}|$. Positive values of *DIFFRET* imply that presentations are more informative than discussions, and negative values imply the opposite.

---

[9] We do not focus on trading volume because disclosures that decrease consensus will increase trading volume (Holthausen and Verrecchia 1990). Thus, if managers provide disclosures that are confusing or ambiguous, then they could potentially increase trading volume (but may not be indicative of greater information content). In addition, prior research suggests that trading volume is exceptionally high for at least several hours (and possibly even days) after the earnings announcement (Lee 1992; Bamber 1987). This fact makes it difficult to disentangle the abnormal trading volume occurring during the presentation/discussion that is the result of information released during the call and the abnormal trading volume related to the earnings release. In contrast, Lee (1992) finds that price adjustments following an earnings announcement occur quickly, generally within the first hour of trading. This makes it less likely that returns during the presentation/discussion are the result of the earnings announcement.

[10] The subsample was gathered for a separate project and represents a random sample of calls made during this time period. We compared this subsample to the overall sample and do not find statistically significant differences between the two in terms of size, analyst following, or any of our performance measures. Thus, we believe the subsample is a representative sample for which to estimate the average words spoken per minute during each segment of the call.

[11] The interquartile range for the delay in the presentation (discussion) is 56 (12) seconds. Starting the computation of returns at the 1st quartile rather than the median would add roughly 28 (6) seconds to the presentation (discussion) return period, which is only 2.5 percent (0.3 percent) of the total time of the presentation (discussion). Thus, we do not believe this assumption significantly affects our results.

Descriptive Statistics

Table 2, Panel A reports descriptive statistics on the length of conference calls. The average length of the presentation (discussion) is 18 (28) minutes, or 2,928 (4,439) words. The interquartile ranges for $\#MIN^{PRES}$ and $\#MIN^{DISC}$ are 10 and 17 minutes, respectively. Thus, discussion periods tend to be longer and are somewhat more variable in length than presentations. If length is related to information content, then this observation suggests that discussion periods are more informative than presentation periods.

We also provide evidence on within-firm variation in call length by computing the interquartile (IQ) range of $\#MIN^{PRES}$ and $\#MIN^{DISC}$ (IQ^{PRES} and IQ^{DISC}) by firm. We note that the average within-firm *IQ* range for discussion periods (10.5 minutes) is over twice as large as the average *IQ* range for presentation periods (4.7 minutes), indicating significantly greater within-firm variation in discussion period length. This difference is not surprising, as firms have greater control over the length of presentations than over discussion periods. In addition, presentations likely follow a general pattern for a given firm, leading to less variation in presentations.

The mean abnormal absolute return during the presentation ($|ABRET^{PRES}|$) is 0.40 percent, while the mean abnormal absolute return during the discussion ($|ABRET^{DISC}|$) is 0.51 percent.[13] Both t-tests as well as signed rank tests indicate that the presentation and discussion period returns are significantly greater than zero ($p < 0.001$), suggesting that both segments of the call are informative to the market. In addition, the average difference in returns (*DIFFRET*) is −0.11 percent, indicating that, on average, discussion periods have greater absolute abnormal returns than presentation periods. Both t-tests and signed rank tests indicate that this value is significantly different from zero ($p < 0.001$). In addition, in 52 percent of firm quarters *DIFFRET* is negative (untabulated), which is significant using a Chi-square test ($\chi^2 = 57.92$; $p < 0.001$).

These results provide initial evidence that both presentation and discussion sessions are incrementally informative to the market. Although prior research suggests that price adjustments generally occur very quickly following an earnings release (Lee 1992), it is possible that the returns during the presentation and discussion period are abnormally high not because of information revealed during the call, but rather because of a delayed reaction to the earnings announcement. Note, however, that to the extent there is a delayed reaction and this reaction dissipates over time, this possibility biases against finding higher information content in the discussion period relative to

---

[12] We use quote data rather than trade data to calculate returns to avoid the bid-ask bounce problem. Our calculation is consistent with Lee (1992) who also uses the quote midpoint to calculate returns and subtracts an average return during a similar time interval over non-announcement days. We use the median rather than the mean because we want to control for time-period-specific price movements that occur on a typical day. However, we also test the sensitivity of our results to using the mean return on non-conference call days and we find inferentially similar results.

[13] Frankel et al. (1999) report average abnormal absolute returns during the entire conference call period of 0.46 percent, which is somewhat smaller than the average abnormal return for the entire call period in our sample (0.40 percent + 0.51 percent = 0.91 percent). However, Frankel et al. (1999) use as their control period the 75 minutes ending 15 minutes prior to the conference call under the assumption that the related earnings release occurs prior to that time (the authors hand-collect the time of the related announcement for their sample of 164 conference calls). If we compare our absolute returns, unadjusted for the control period, our returns are similar: 1.34 percent in our sample versus 1.23 percent in their sample.

TABLE 2

Descriptive Statistics on Dependent Variables

Panel A: Descriptive Statistics on Call Length (n = 10,062)

Variable	Mean	1st Percentile	25th Percentile	Median	75th Percentile	99th Percentile	Std. Dev.	IQ Range
$\#MIN^{PRES}$	18.29	3.78	12.79	17.48	22.73	41.18	7.79	9.93
$\#MIN^{DISC}$	28.25	4.47	18.95	27.64	36.20	65.08	13.09	17.25
$LENGTH^{PRES}$	2928.17	605.00	2047.00	2798.00	3638.00	6591.00	1246.72	1591.00
$LENGTH^{DISC}$	4438.53	702.00	2977.00	4342.00	5687.00	10223.00	2055.93	2710.00
IQ^{PRES}	4.65	0.69	2.50	3.93	6.02	15.95	3.32	
IQ^{DISC}	10.46	1.56	6.59	9.50	13.43	26.97	5.42	

Panel B: Descriptive Statistics on Abnormal Absolute Returns (n = 10,062)

Variable	Mean	1st Percentile	25th Percentile	Median	75th Percentile	99th Percentile	Std. Dev.	t-stat p-value	Signed Rank p-value
$\lvert ABRET^{PRES}\rvert$	0.0040	−0.0037	−0.0004	0.0014	0.0054	0.0388	0.0083	0.001	0.001
$\lvert ABRET^{DISC}\rvert$	0.0051	−0.0036	−0.0003	0.0019	0.0067	0.0470	0.0099	0.001	0.001
$DIFFRET$	−0.0011	−0.0401	−0.0041	−0.0002	0.0028	0.0323	0.0113	0.001	0.001

Panel C: Correlation Matrix[a]

Variable	$\lvert ABRET^{PRES}\rvert$	$\lvert ABRET^{DISC}\rvert$	$\lvert RET^{DAYB4}\rvert$	$LENGTH^{PRES}$	$LENGTH^{DISC}$
$\lvert ABRET^{PRES}\rvert$	1.000	**0.193**	**0.163**	**0.105**	**0.058**
$\lvert ABRET^{DISC}\rvert$	**0.245**	1.000	**0.182**	0.001	**0.145**
$\lvert RET^{DAYB4}\rvert$	**0.230**	**0.225**	1.000	−0.012	**0.071**
$LENGTH^{PRES}$	**0.125**	0.003	−0.016	1.000	**0.111**
$LENGTH^{DISC}$	**0.033**	**0.161**	**0.061**	**0.080**	1.000

---

[a] Spearman correlations shown above the diagonal; Pearson correlations shown below the diagonal. Correlations in bold are significant at less than 0.001, the remaining correlations are not significant at the 10 percent level.

Variable Definitions:

$\#MIN^{PRES}$, $\#MIN^{DISC}$ = estimated number of minutes of the presentation and discussion, based on the median number of words spoken per minute based on a subsample of 1,263 calls for which we have the exact start and end times of the presentation and discussion;

$LENGTH^{PRES}$, $LENGTH^{DISC}$ = number of words spoken during the presentation and discussion, respectively;

IQ^{PRES}, IQ^{DISC} = firm-specific interquartile ranges of the number of minutes in the presentation and discussion, respectively;

$\lvert RET^{PRES}\rvert$ ($\lvert RET^{DISC}\rvert$) = absolute value of the difference between the quote midpoint at the start of the presentation (discussion) and the quote midpoint at the end of the presentation (discussion), scaled by the quote midpoint at the start of the presentation (discussion);

$\lvert ABRET^{PRES}\rvert$ ($\lvert ABRET^{DISC}\rvert$) = $\lvert RET^{PRES}\rvert$ ($\lvert RET^{DISC}\rvert$) less the median value of all absolute returns measured during the same time period on non-conference call days during the quarter;

$DIFFRET$ = $\lvert ABRET^{PRES}\rvert$ less $\lvert ABRET^{DISC}\rvert$; and

$\lvert RET^{DAYB4}\rvert$ = absolute value of the quote midpoint at the start of the conference call less the quote midpoint at the same time one trading day before the conference call ($MID^{dayprior}$) divided by $MID^{dayprior}$ (from the TAQ database).

the presentation period. Nevertheless, to provide evidence on this possibility, we examine the relation between presentation and discussion period abnormal returns and the absolute returns in the 24 hours prior to the conference call ($\lvert RET^{DAYB4}\rvert$). Since the vast majority of earnings press releases are released shortly before the conference call, this variable should capture the news in the earnings press release.[14] Specifically, we take the quote midpoint at the start of the conference call (MID^{START}) less the quote midpoint at the same time one trading day before the conference call (MID^{DAYB4}) divided by MID^{DAYB4}.[15] To the extent that delayed reactions are greater when news in the earnings announcement is greater, there should be a positive association between $|RET^{DAYB4}|$ and $|ABRET^{PRES}|$ and $|ABRET^{DISC}|$.

Table 2, Panel C presents a correlation matrix of the information content variables (as well as measures of call length). Spearman (Pearson) correlations are reported above (below) the diagonal. Both $|ABRET^{PRES}|$ and $|ABRET^{DISC}|$ are highly correlated with $|RET^{DAYB4}|$ ($r = 0.225$ and $r = 0.245$ for the presentation and discussion, respectively), suggesting the possibility of a delayed reaction to news in the earnings announcement. However, it is also possible that managers provide more disclosures during conference calls when there is more news in the earnings announcement. The fact that the length of the discussion ($LENGTH^{DISC}$) is also highly correlated with $|RET^{DAYB4}|$ suggests this possibility. Nevertheless, it is important that we control for the possibility of delayed reactions when analyzing the information content of the presentation and discussion. In addition, we note that $|ABRET^{PRES}|$ and $|ABRET^{DISC}|$ are highly correlated with each other ($r = 0.24$), suggesting that returns during the discussion period also may be a delayed reaction to the presentation.[16] Thus, in subsequent tests, we control for the absolute returns during the presentation when analyzing the absolute returns during the discussion.[17]

IV. EMPIRICAL DESIGN AND RESULTS

Information Content of Presentation and Discussion Sessions

Our formal tests of the information content of the presentation and discussion sessions are based on the following regressions:

$$|ABRET^{PRES}| = \beta_0 + \beta_1 |RET^{DAYB4}| + \varepsilon; \tag{1}$$

$$|ABRET^{DISC}| = \beta_0 + \beta_1 |RET^{DAYB4}| + \beta_2 |ABRET^{PRES}| + \varepsilon. \tag{2}$$

---

[14] For approximately 98 percent of our sample, the conference call date coincides with the earnings announcement date per Compustat or is one day later. Thus, returns in the 24 hours prior to the call capture the market reaction to the press release in the vast majority of cases.

[15] Note that we do not adjust the returns for the median return on non-conference call days (as we do with $|ABRET^{PRES}|$ and $|ABRET^{DISC}|$) because intra-day trading patterns are not a concern with the variable since it is measured over an entire trading day.

[16] We also examined the correlation between *signed* returns during the earnings release and *signed* returns during the presentation/discussion. To the extent there is a delayed price reaction, one would expect the signed returns to be correlated. The correlation between earnings announcement signed returns and (1) presentation period signed returns and (2) discussion period signed returns are both insignificant. However, we do note a significantly positive correlation between signed returns during the presentation and signed returns during the discussion ($r = 0.049$), suggesting the possibility of a delayed reaction to the presentation occurring during the discussion.

[17] We also note that the length of the presentation (discussion) is significantly correlated with the information content of the presentation (discussion), indicating that longer presentations/discussions have greater information content. Note, however, that the correlations are far from unity ($r = 0.125$ and 0.16 for presentations and discussions, respectively), suggesting that these potentially measure different aspects of information content. In subsequent sensitivity tests, we use length as an alternative measure of information content and find similar results. Our results, however, do not speak to whether the information *per words spoken* differs between the presentation and discussion. Such an analysis is problematic because presentations are prepared in advance, whereas discussions are based on spontaneous speech, making it difficult to compare the "efficiency" of words spoken in communicating information.

The intercept (β_0) in each equation represents the average abnormal absolute return during each segment of the call, controlling for the information content of prior news events. In the case of the presentation (Equation (1)), we control for the information content of the earnings announcement press release ($|RET^{DAYB4}|$). In the case of the discussion (Equation (2)), we control for both the information content of the press release as well as the information content of the presentation portion of the call. Because it is possible that the magnitude of returns that occurs during the presentation and discussion segments of the call has a firm-specific component, we calculate standard errors clustered by firm (Petersen 2009).[18] In addition, to ensure our results are not driven by extreme observations, we winsorize all variables at the 1 percent level.

The results of this analysis are presented in Table 3, Panel A. The first (second) set of columns presents the results for the returns during the presentation (discussion). Consistent with our univariate correlations, the coefficient on $|RET^{DAYB4}|$ is positive and statistically significant in both sets of regressions, as is the coefficient on $|ABRET^{PRES}|$ in the regression of discussion period returns. More importantly, the intercepts in both equations are positive and statistically significant, indicating that both the presentation and discussion segments have incremental information content over the accompanying press release and, in the case of the discussion, over information released during the presentation. The value of the intercept indicates that the average abnormal absolute return, after controlling for the information content of prior news events, is roughly 0.2 percent for both the presentation and discussion.[19]

We also examine differences in information content between the two segments by running a regression similar to Equation (1) but with *DIFFRET* as the dependent variable. The results of this analysis are presented in Table 3, Panel B. The coefficient on $|RET^{DAYB4}|$ is negative and significant, but only at the 10 percent level.[20] More importantly, the intercept is negative and statistically significant, consistent with the discussion segment of the call having greater information content than the presentation. The magnitude of the intercept indicates that, on average, abnormal absolute returns during the discussion session are 0.08 percent higher than absolute returns during the presentation. The magnitude of the difference appears small, but represents a 34 percent increase over the average abnormal absolute returns during the presentation (0.08 percent ÷ 0.23 percent).

The fact that the discussion portion of the call is relatively more informative than the presentation suggests that analyst participation is an important part of the incremental informativeness of conference calls. To further investigate the role of analysts in the information content of the presentation and discussion, we calculate the quintile rank (by year) of the number of

---

[18] For example, sophisticated investors may be more likely to listen to conference calls and trade during the call period, potentially impacting the speed with which information is impounded in price. We do not use two-way clustered standard errors (as suggested by Gow et al. [2010]) because the probability of cross-sectional dependence in our dependent variable is low (i.e., the likelihood of significant overlap in the exact timing of presentation and discussion periods in our sample is low given the narrow windows used to measure the returns).

[19] The fact that average returns in this analysis have roughly equivalent magnitudes might suggest that both segments have roughly the same information content. However, if we exclude presentation returns from our discussion period regression, we find average returns during the discussion of 0.3 percent, which are roughly 30 percent larger in magnitude than presentation period returns. While including presentation returns in the discussion period regression alleviates concerns of delayed price reactions, it also reduces the impact of any discussion period disclosures that are correlated with presentation disclosures.

[20] Note that if there was a delayed reaction to news in the earnings announcement, then one would expect a *positive* (not negative) coefficient on $|RET^{DAYB4}|$ because presentation abnormal returns would be greater than discussion period abnormal returns (since presentations occur before discussions). It is possible that discussion period returns are affected by information in the presentation; however, including $|ABRET^{PRES}|$ in this regression is problematic because $|ABRET^{PRES}|$ is part of the definition of *DIFFRET*. Nevertheless, in untabulated analysis, we include $|ABRET^{PRES}|$ in the regression and continue to find a significantly negative intercept. Not surprisingly, the coefficient on $|ABRET^{PRES}|$ is positive and extremely significant.

TABLE 3

Analysis of the Information Content of Presentation and Discussion Sessions

Panel A: Incremental Information Content of Presentation and Discussion Sessions[a]

		$\lvert ABRET^{PRES}\rvert$			$\lvert ABRET^{DISC}\rvert$		
Variable[c]	**Pred. Sign**	**Coeff.**	**t-stat**	**p-value**	**Coeff.**	**t-stat**	**p-value**
Intercept	+	0.0023	20.99	0.001	0.0021	15.15	0.001
$\lvert RET^{DAYB4}\rvert$	+	0.0481	15.31	0.001	0.0444	12.84	0.001
$\lvert ABRET^{PRES}\rvert$	+				0.2315	13.71	0.001
			n = 10,062, $R^2 = 0.055$			n = 10,062, $R^2 = 0.091$	

Panel B: Relative Incremental Information Content of Presentation versus Discussion[a]

			DIFFRET	
Variable[c]	**Pred. Sign**	**Coeff.**	**t-stat**	**p-value**
Intercept	–	−0.0008	−5.41	0.001
$\lvert RET^{DAYB4}\rvert$	?	−0.0082	−1.87	0.062
			n = 10,062, $R^2 = 0.003$	

Panel C: Effect of Analyst Following on Incremental Information Content[b]

		$\lvert ABRET^{PRES}\rvert$			$\lvert ABRET^{DISC}\rvert$		
Variable[c]	**Pred. Sign**	**Coeff.**	**t-stat**	**p-value**	**Coeff.**	**t-stat**	**p-value**
Intercept	+	0.0023	13.25	0.001	0.0015	8.26	0.001
$\lvert RET^{DAYB4}\rvert$	+	0.0480	15.38	0.001	0.0452	13.14	0.001
$\lvert ABRET^{PRES}\rvert$	+				0.2313	13.67	0.001
R_AF	+	0.0000	0.33	0.628	0.0003	4.24	0.001
			n = 10,062, $R^2 = 0.055$			n = 10,062, $R^2 = 0.094$	

Panel D: Effect of Analyst Following on Relative Incremental Information Content[b]

		DIFFRET			*DIFFLENGTH*		
Variable[c]	**Pred. Sign**	**Coeff.**	**t-stat**	**p-value**	**Coeff.**	**t-stat**	**p-value**
Intercept	+	0.0002	−0.76	0.776	−0.4044	−5.66	0.001
$\lvert RET^{DAYB4}\rvert$	?	−0.0091	−2.06	0.039	−5.7267	−7.02	0.001
R_AF	–	−0.0003	−3.79	0.001	−0.4571	−14.02	0.001
			n = 10,062, $R^2 = 0.003$			n = 10,062, $R^2 = 0.090$	

---

[a] Panels A and B present results of regressions of (1) abnormal absolute returns during the presentation ($\lvert ABRET^{PRES}\rvert$) and discussion ($\lvert ABRET^{DISC}\rvert$) and (2) the difference in abnormal returns between the presentation and discussion (*DIFFRET*) on absolute returns during the earnings announcement period ($\lvert RET^{DAYB4}\rvert$) and, in the case of returns during the discussion, on the absolute returns during the presentation. The intercept in each regression represents the mean abnormal absolute returns, controlling for the information content of prior events, during each segment of the call or, in the case of *DIFFRET*, the mean difference in absolute abnormal returns (where positive [negative] values indicate the presentation [discussion] has greater information content). Standard errors are clustered by firm.

[b] Panels C and D present results of regressions similar to Panels A and B but including the within-year quintile rank of the firm's analyst following. A positive coefficient on *R_AF* in Panel C indicates greater information content for firms with higher analyst following. A negative coefficient on *R_AF* in Panel D indicates the relative information content of the discussion over the presentation is greater for firms with higher analyst following. In addition, we run similar regressions using the difference in the length (in number of words) of the presentation and discussion sessions (*DIFFLENGTH*). Standard errors are clustered by firm.

[c] See Table 2 for variable definitions.

analysts issuing a forecast for the quarter (*R_AF*) and include this variable in our prior regressions.[21]

The results of these analyses are presented in Table 3, Panels C and D. For presentation returns, the coefficient on *R_AF* is not statistically different from zero, indicating that greater analyst following does not increase the information content of management presentations. In contrast, the abnormal absolute returns during the discussion portion of the call increase as analyst following increases, which is more consistent with the notion that analysts' active involvement in the discussion session is the source of the greater information content of the discussion session (as opposed to a delayed price reaction to the presentation).

Also, as shown in Table 3, Panel D, the difference in returns between the presentation and discussion is negatively related to analyst following. The intercept is no longer statistically significant, indicating that firms in the lowest analyst following quintile have equally informative presentation and discussion sessions. In contrast, the information content of the discussion is 0.14 percent higher than the presentation for firms in the highest analyst following quintile $((4 \times 0.0003 = 0.0012) + 0.0002 = 0.0014)$, a difference that represents a 38 percent increase over the average abnormal returns during the presentation session for firms in this quintile.[22]

As additional evidence that the greater information content of the discussion session (relative to the presentation) is due to analyst involvement, we also examine the difference in the relative length (in words) of each segment. *DIFFLENGTH* is defined as the difference between $LENGTH^{PRES}$ and $LENGTH^{DISC}$. Results are reported in the last set of columns in Table 3, Panel D. As with absolute returns, we find that the coefficient on *R_AF* is significantly negative, indicating that discussion periods are longer relative to presentations when analyst following is higher.[23]

Overall, these results support our conjecture that, while both presentations and discussion sessions have incremental information content over the accompanying earnings press release, discussion sessions tend to provide greater relative information content likely as a result of analysts' active involvement in these question-and-answer sessions.

The Relation between Information Content and Firm Performance

We next examine whether managers' incentives to disclose information varies based on firm performance. We use two accounting and two market-based measures of firm performance. The accounting-based measures are (1) return on assets for the quarter (*ROA*) and (2) an indicator variable equal to 1 if the firm failed to meet analysts' consensus forecasts for the quarter (*MISS*). The market-based measures are (1) market-adjusted returns during the quarter, excluding the earnings announcement (RET^{QTR}) and (2) the returns during the 24 hours prior to the conference call (RET^{DAYB4}). Note that for each set of measures, one captures performance over the entire quarter, while the other captures the unexpected performance at the time of the earnings release. Exact variable definitions are provided in Table 4.

Because we are specifically interested in the impact of performance on the information content of presentations and discussion sessions, it is important to control for potential correlated omitted

---

[21] Alternatively, we measured the number of analysts who participate in each conference call held by the firm and find similar results.

[22] The mean of *DIFFRET* for quintiles one to five, respectively, are −0.0002, −0.0009, −0.0013, −0.0013, −0.0014. Thus, the difference in abnormal absolute returns decreases nearly monotonically across the five quintile groups.

[23] Note, however, that the intercept in this regression is negative and significant, indicating that even for the lowest quintile of analyst following, discussions are significantly longer than presentations. In addition, in untabulated analysis we also find that the length of the presentation is positively related to analyst following, in contrast to absolute returns during the presentation, which are unrelated to analyst following. Thus, firms with greater analyst following provide longer presentations but these presentations are not necessarily more informative.

TABLE 4

Firm Performance and the Information Content of Presentations and Discussions

Panel A: Univariate Statistics

Variable	Mean	1st Quartile	Median	3rd Quartile	Std. Deviation
ROA	0.008	0.003	0.010	0.020	0.043
RET^{QTR}	−0.008	−0.102	−0.016	0.075	0.170
RET^{DAYB4}	0.001	−0.020	0.001	0.023	0.051
$MISS$	0.285	0.000	0.000	1.000	0.451
$\lvert RET^{QTR}\rvert$	0.121	0.043	0.090	0.162	0.120
$\lvert RET^{DAYB4}\rvert$	0.034	0.009	0.022	0.045	0.038
$\lvert FE\rvert$	0.005	0.000	0.001	0.003	0.024
$\lvert SPITEM\rvert$	0.003	0.000	0.000	0.001	0.031
$\lvert DISOP\rvert$	0.001	0.000	0.000	0.000	0.007
$M\&A^{CC}$	0.193	0.000	0.096	0.295	0.241
$RESTR^{CC}$	0.044	0.000	0.000	0.038	0.106
$\Delta SALES$	0.026	0.001	0.015	0.041	0.066
AF	7.350	3.000	6.000	10.000	5.888
LNA	7.277	6.053	7.190	8.371	1.782

Panel B: Correlation Matrix[a]

Variable	ROA	RET^{QTR}	RET^{DAYB4}	$MISS$	$\lvert RET^{QTR}\rvert$	$\lvert RET^{DAYB4}\rvert$	$\lvert FE\rvert$	$\lvert SPITEM\rvert$	$\lvert DISOP\rvert$	$M\&A^{CC}$	$RESTR^{CC}$	$\Delta SALES$	AF	LNA
ROA	1.00	0.05	0.09	−0.18	−0.03	0.02	−0.05	−0.22	−0.02	−0.01	−0.08	0.24	0.07	0.08
RET^{QTR}	−0.01	1.00	−0.05	−0.07	−0.07	−0.01	−0.04	−0.01	0.01	0.00	−0.01	0.05	0.00	0.07
RET^{DAYB4}	0.03	−0.07	1.00	−0.27	0.00	0.03	0.05	−0.01	0.00	0.01	0.00	0.09	−0.04	−0.01
$MISS$	−0.08	−0.08	−0.28	1.00	0.01	0.05	0.17	0.03	0.00	0.00	0.01	−0.10	0.02	0.03
$\lvert RET^{QTR}\rvert$	−0.01	0.07	0.00	0.00	1.00	0.04	0.04	0.03	0.00	0.01	0.03	−0.02	−0.06	−0.08
$\lvert RET^{DAYB4}\rvert$	0.04	−0.02	−0.05	0.07	0.05	1.00	0.12	0.03	0.00	−0.02	0.00	0.01	0.02	0.04
$\lvert FE\rvert$	−0.04	−0.03	0.00	0.06	0.03	0.04	1.00	0.09	0.08	−0.01	0.04	−0.01	−0.04	−0.03
$\lvert SPITEM\rvert$	−0.67	0.02	0.00	0.01	0.01	−0.01	0.03	1.00	0.07	0.02	0.15	−0.02	−0.01	−0.01
$\lvert DISOP\rvert$	0.00	0.03	0.00	0.01	0.02	−0.01	0.02	0.01	1.00	−0.01	0.05	−0.03	−0.01	0.02
$M\&A^{CC}$	0.00	0.01	0.02	0.00	−0.01	−0.02	0.00	0.00	−0.01	1.00	−0.02	0.08	−0.01	0.06

(continued on next page)

TABLE 4 (continued)

Variable	ROA	RET^{QTR}	RET^{DAYB4}	$MISS$	$\lvert RET^{QTR}\rvert$	$\lvert RET^{DAYB4}\rvert$	$\lvert FE\rvert$	$\lvert SPITEM\rvert$	$\lvert DISOP\rvert$	$M\&A^{CC}$	$RESTR^{CC}$	$\Delta SALES$	AF	LNA
$RESTR^{CC}$	−0.02	0.00	0.02	0.01	0.02	−0.01	0.05	0.01	0.02	−0.04	1.00	−0.06	−0.03	−0.01
$\Delta SALES$	0.10	0.01	0.06	−0.07	0.01	0.02	0.01	−0.02	−0.01	0.05	−0.03	1.00	0.02	0.10
AF	0.02	−0.02	−0.03	0.02	−0.07	0.02	−0.01	0.01	−0.02	−0.01	−0.02	−0.02	1.00	0.27
LNA	0.02	0.00	−0.02	0.03	−0.06	0.04	−0.01	0.01	0.00	0.07	−0.01	0.04	0.26	1.00

Panel C: Relation between Firm Performance and Abnormal Absolute Returns during the Presentation[b]

		$\lvert ABRET^{PRES}\rvert$		$\lvert ABRET^{PRES}\rvert$		$\lvert ABRET^{PRES}\rvert$		$\lvert ABRET^{PRES}\rvert$		$\lvert ABRET^{PRES}\rvert$		$\lvert ABRET^{PRES}\rvert$	
Variable	**Pred. Sign**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**
ROA	+/−	−0.0023	−0.73	−0.0019	−0.37					−0.0023	−0.52	−0.0009	−0.19
$MISS$	+/−	**0.0007**	**4.24**	**0.0005**	**2.69**					*0.0003*	*1.85*	0.0001	0.44
RET^{QTR}	+/−					*−0.0009*	*−1.89*	−0.0006	−1.29	*−0.0008*	*−1.69*	−0.0006	−1.24
RET^{DAYB4}	+/−					**−0.0130**	**−8.50**	**−0.0122**	**−8.02**	**−0.0121**	**−7.55**	**−0.0120**	**−7.55**
$\lvert RET^{QTR}\rvert$	+/−			0.0004	0.54			0.0004	0.56			0.0004	0.56
$\lvert RET^{DAYB4}\rvert$	+/−			**0.0310**	**13.73**			**0.0307**	**13.69**			**0.0307**	**13.64**
$\lvert FE\rvert$	+/−			**0.0329**	**3.01**			**0.0362**	**3.37**			**0.0353**	**3.24**
$\lvert SPITEM\rvert$	+/−			*0.0179*	*1.63*			**0.0201**	**1.99**			*0.0193*	*1.76*
$\lvert DISOP\rvert$	+/−			**0.0658**	**1.99**			**0.0644**	**1.95**			**0.0646**	**1.96**
$M\&A^{CC}$	+/−			0.0201	0.41			0.0287	0.59			0.0283	0.58
$RESTR^{CC}$	+/−			0.0179	0.18			0.0340	0.34			0.0326	0.33
$\Delta SALES$	+/−			−0.0004	−0.19			0.0005	0.21			0.0006	0.28
AF	+/−			0.0001	1.39			0.0001	1.16			0.0001	1.17
LNA	+/−			**0.0019**	**3.88**			**0.0018**	**3.80**			**0.0018**	**3.79**
$Q2$	+/−	0.0001	0.33	−0.0001	−0.65	0.0000	0.14	−0.0001	−0.78	0.0000	0.21	−0.0001	−0.76
$Q3$	+/−	−0.0002	−0.79	*−0.0003*	*−1.77*	−0.0002	−1.04	**−0.0004**	**−1.95**	−0.0002	−1.01	**−0.0004**	**−1.93**
$Q4$	+/−	0.0001	0.50	0.0000	−0.09	0.0000	0.25	−0.0001	−0.35	0.0001	0.27	−0.0001	−0.33
$Y04$	+/−	**0.0004**	**2.42**	0.0003	1.52	**0.0005**	**2.66**	*0.0003*	*1.76*	**0.0005**	**2.68**	*0.0003*	*1.76*
$Y05$	+/−	0.0002	1.02	−0.0002	−1.19	0.0002	1.26	−0.0002	−0.94	0.0002	1.24	−0.0002	0.94
Firm Fixed Effects		Included		Included		Included		Included		Included		Included	

(continued on next page)

TABLE 4 (continued)

Panel D: Relation between Firm Performance and Abnormal Absolute Returns during the Discussion[c]

Variable	Pred. Sign	$\lvert ABRET^{DISC}\rvert$		$\lvert ABRET^{DISC}\rvert$		$\lvert ABRET^{DISC}\rvert$		$\lvert ABRET^{DISC}\rvert$		$\lvert ABRET^{DISC}\rvert$		$\lvert ABRET^{DISC}\rvert$	
		Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
ROA	−	0.0023	0.43	−0.0015	−0.25					0.0035	0.65	−0.0002	−0.04
MISS	+	**0.0008**	**4.07**	**0.0007**	**3.34**					*0.0003*	*1.51*	0.0002	0.78
RET^{QTR}	−					**−0.0017**	**−3.06**	**−0.0016**	**−2.86**	**−0.0017**	**−2.91**	**−0.0016**	**−2.77**
RET^{DAYB4}	−					**−0.0160**	**−8.75**	**−0.0161**	**−8.85**	**−0.0153**	**−7.98**	**−0.0157**	**−8.25**
$\lvert RET^{QTR}\rvert$	+/−			**0.0022**	**2.43**			**0.0022**	**2.49**			**0.0022**	**2.49**
$\lvert RET^{DAYB4}\rvert$	+/−			**0.0295**	**10.81**			**0.0295**	**10.86**			**0.0294**	**10.80**
$\lvert FE\rvert$	+/−			0.0093	0.71			0.0138	1.07			0.0122	0.93
$\lvert SPITEM\rvert$	+/−			0.0124	0.94			0.0154	1.27			0.0151	1.15
$\lvert DISOP\rvert$	+/−			0.0364	0.92			0.0359	0.91			0.0362	0.92
$M\&A^{CC}$	+/−			−0.0003	−0.52			−0.0002	−0.32			−0.0002	−0.32
$RESTR^{CC}$	+/−			*−0.0022*	*−1.84*			*−0.0020*	*−1.67*			*−0.0020*	*−1.68*
$\Delta SALES$	+/−			**0.0069**	**2.52**			**0.0081**	**3.06**			**0.0083**	**3.03**
AF	+/−			0.0000	−0.86			−0.0001	−1.17			−0.0001	−1.17
LNA	+/−			**0.0016**	**2.71**			**0.0015**	**2.58**			**0.0015**	**2.57**
Q2	+/−	**0.0007**	**3.09**	**0.0004**	**2.52**	**0.0007**	**2.94**	**0.0005**	**2.35**	**0.0007**	**2.94**	**0.0005**	**2.36**
Q3	+/−	**0.0005**	**2.20**	*0.0004*	*1.78*	*0.0004*	*1.93*	0.0004	1.52	*0.0004*	*1.90*	0.0004	1.53
Q4	+/−	0.0003	1.30	0.0003	1.31	0.0002	0.95	0.0002	0.89	0.0002	0.96	0.0002	0.91
Y04	+/−	*0.0004*	*1.73*	0.0003	1.39	**0.0005**	**2.23**	*0.0004*	*1.89*	**0.0005**	**2.18**	*0.0004*	*1.87*
Y05	+/−	**0.0005**	**2.10**	0.0002	0.92	**0.0006**	**2.59**	0.0004	1.40	**0.0005**	**2.48**	0.0003	1.37
$\lvert ABRET^{PRES}\rvert$	+	**0.1832**	**14.37**	**0.1594**	**12.44**	**0.1750**	**13.73**	**0.1508**	**11.77**	**0.1747**	**13.71**	**0.1508**	**11.77**
Firm Fixed Effects		Included		Included		Included		Included		Included		Included	

Panel E: Relation between Firm Performance and Relative Information Content of Presentation versus Discussion[d]

Variable	Pred. Sign	*DIFFRET*		*DIFFRET*		*DIFFRET*		*DIFFRET*		*DIFFRET*		*DIFFRET*	
		Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
ROA	+	−0.0050	−0.77	−0.0001	−0.01					−0.0054	−0.83	−0.0006	−0.08
MISS	−	−0.0002	−0.96	−0.0003	−1.18					0.0000	−0.19	−0.0001	−0.38

(continued on next page)

TABLE 4 (continued)

		DIFFRET		*DIFFRET*		*DIFFRET*		*DIFFRET*		*DIFFRET*		*DIFFRET*	
Variable	**Pred. Sign**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**
RET^{QTR}	+					*0.0010*	*1.44*	*0.0011*	*1.58*	*0.0010*	*1.42*	*0.0011*	*1.54*
RET^{DAYB4}	+					**0.0053**	**2.38**	**0.0058**	**2.60**	**0.0053**	**2.28**	**0.0055**	**2.38**
$\lvert RET^{QTR}\rvert$	+/−			*−0.0018*	*−1.68*			*−0.0018*	*−1.71*			*−0.0018*	*−1.71*
$\lvert RET^{DAYB4}\rvert$	+/−			−0.0035	−1.06			−0.0034	−1.03			−0.0033	−1.00
$\lvert FE\rvert$	+/−			0.0183	1.15			0.0169	1.07			0.0178	1.11
$\lvert SPITEM\rvert$	+/−			0.0026	0.16			0.0017	0.11			0.0012	0.08
$\lvert DISOP\rvert$	+/−			0.0189	0.39			0.0188	0.39			0.0186	0.39
$M\&A^{CC}$	+/−			0.0475	0.67			0.0427	0.60			0.0429	0.60
$RESTR^{CC}$	+/−			0.2369	1.61			0.2288	1.55			0.2295	1.56
$\Delta SALES$	+/−			**−0.0072**	**−2.17**			**−0.0077**	**−2.37**			**−0.0077**	**−2.31**
AF	+/−			0.0001	1.50			*0.0001*	*1.63*			*0.0001*	*1.63*
LNA	+/−			0.0000	0.01			0.0001	0.09			0.0001	0.09
$Q2$	+/−	**−0.0007**	**−2.36**	**−0.0007**	**−2.43**	**−0.0007**	**−2.34**	**−0.0007**	**−2.37**	**−0.0006**	**−2.30**	**−0.0007**	**−2.37**
$Q3$	+/−	**−0.0006**	**−2.26**	**−0.0007**	**−2.47**	**−0.0006**	**−2.18**	**−0.0007**	**−2.37**	**−0.0006**	**−2.14**	**−0.0007**	**−2.37**
$Q4$	+/−	−0.0002	0.79	−0.0003	−1.12	−0.0002	−0.64	−0.0003	−0.93	−0.0002	−0.64	−0.0003	−0.93
$Y04$	+/−	0.0000	−0.06	−0.0001	−0.26	−0.0001	−0.32	−0.0001	−0.52	−0.0001	−0.26	−0.0001	−0.51
$Y05$	+/−	−0.0003	−1.15	−0.0004	−1.44	−0.0004	−1.41	*−0.0005*	*−1.68*	−0.0004	−1.34	*−0.0005*	*−1.67*
	Firm Fixed Effects	Included		Included		Included		Included		Included		Included	

Panel F: Relation between Firm Performance and Length of Presentation and Discussion[e]

	$LENGTH^{PRES}$			$LENGTH^{DISC}$			$DIFFLENGTH$		
Variable	**Pred.**	**Coeff.**	**t-stat**	**Pred.**	**Coeff.**	**t-stat**	**Pred.**	**Coeff.**	**t-stat**
ROA	+/−	0.3438	0.75	−	0.5740	0.59	+	−0.2302	−0.21
$MISS$	+/−	**0.0377**	**2.24**	+	**0.0682**	**1.92**	−	−0.0305	−0.76
RET^{QTR}	+/−	**−0.1761**	**−4.01**	−	**−0.4832**	**−5.23**	+	**0.3071**	**2.95**
RET^{DAYB4}	+/−	**−0.3781**	**−2.57**	−	**−2.0982**	**−6.78**	+	**1.7202**	**4.95**
$\lvert RET^{QTR}\rvert$	+/−	−0.0374	−0.55	+/−	0.2227	1.55	+/−	−0.2602	−1.61
$\lvert RET^{DAYB4}\rvert$	+/−	**0.4544**	**2.18**	+/−	**3.7947**	**8.65**	+/−	**−3.3403**	**−6.77**
$\lvert FE\rvert$	+/−	*1.8705*	*1.85*	+/−	2.5488	1.20	+/−	−0.6783	−0.28

(continued on next page)

TABLE 4 (continued)

	$LENGTH^{PRES}$			$LENGTH^{DISC}$			$DIFFLENGTH$		
Variable	**Pred.**	**Coeff.**	**t-stat**	**Pred.**	**Coeff.**	**t-stat**	**Pred.**	**Coeff.**	**t-stat**
$\|SPITEM\|$	+/–	**5.0412**	**4.96**	+/–	0.1351	0.06	+/–	**4.9061**	**2.04**
$\|DISOP\|$	+/–	–2.3761	–0.78	+/–	–2.1618	–0.34	+/–	–0.2143	0.39
$M\&A^{CC}$	+/–	7.1223	1.58	+/–	*–17.5129*	*–1.85*	+/–	**24.6351**	**2.31**
$RESTR^{CC}$	+/–	**26.8808**	**2.89**	+/–	–9.8172	–0.50	+/–	*36.6980*	*1.67*
$\Delta SALES$	+/–	–0.1316	–0.62	+/–	0.3730	0.84	+/–	–0.5046	–1.01
AF	+/–	–0.0011	–0.25	+/–	**0.1119**	**12.22**	+/–	**–0.1130**	**–10.97**
LNA	+/–	**0.1390**	**3.15**	+/–	**0.2392**	**2.57**	+/–	–0.1002	–0.96
$Q2$	+/–	**0.1201**	**6.80**	+/–	**0.0951**	**2.56**	+/–	0.0250	0.60
$Q3$	+/–	**0.1434**	**7.95**	+/–	**0.0777**	**2.05**	+/–	0.0657	1.54
$Q4$	+/–	**0.3854**	**21.32**	+/–	**0.1288**	**3.38**	+/–	**0.2567**	**6.00**
$Y04$	+/–	**0.0995**	**5.88**	+/–	0.0323	0.91	+/–	*0.0672*	*1.68*
$Y05$	+/–	**0.0533**	**2.76**	+/–	*0.0782*	*1.90*	+/–	–0.0249	–0.54
Firm Fixed Effects		Included			Included			Included	

---

[a] Spearman correlations shown above the diagonal; Pearson correlations shown below the diagonal. For purposes of this table, we subtract the firm-mean from each variable (because subsequent analyses are based on firm fixed effect models).

[b] $|ABRET^{PRES}|$ is the abnormal absolute returns during the presentation portion of the conference call as described in Table 2. *Q2*, *Q3*, and *Q4* are indicator variables equal to 1 if the conference call relates to fiscal quarter 2, 3, and 4, respectively. *Y04* and *Y05* are indicator variables equal to 1 if the conference call is held in 2004 and 2005, respectively. Bolded (italicized) values are significant at the 5 percent (10 percent) level, using a two-tailed test.

[c] $|ABRET^{DISC}|$ is the abnormal absolute returns during the discussion portion of the conference call as described in Table 2. Bolded (italicized) values are significant at the 5 percent (10 percent) level using a one-tailed (two-tailed) test for variables with (without) directional predictions.

[d] *DIFFRET* is the abnormal absolute returns during the presentation less the abnormal absolute returns during the discussion as described in Table 2. Bolded (italicized) values are significant at the 5 percent (10 percent) level using a one-tailed (two-tailed) test for variables with (without) directional predictions.

[e] $LENGTH^{PRES}$ ($LENGTH^{DISC}$) is the length of the presentation (discussion) in words (divided to facilitate reporting of coefficients) as described in Table 2. *DIFFLENGTH* is the length of the presentation less the length of the discussion. Bolded (italicized) values are significant at the 5 percent (10 percent) level using a one-tailed (two-tailed) test for variables with (without) directional predictions.

Variable Definitions:
ROA = earnings before extraordinary items (Compustat quarterly item 8) divided by ending total assets (Compustat quarterly item 44);
RET^{QTR} = market-adjusted returns cumulated from day –92 to day –2 relative to the conference call date;
RET^{DAYB4} = quote midpoint at the start of the conference call (MID^{start}) less the quote midpoint at the same time one trading day before the conference call ($MID^{dayprior}$) divided by $MID^{dayprior}$ (from the TAQ database);
MISS = indicator variable that is equal to 1 if actual EPS for the quarter (per I/B/E/S unsplit-adjusted actuals file) is less than the mean consensus forecasted EPS for the quarter (per I/B/E/S unsplit-adjusted summary file) and 0 otherwise;

(continued on next page)

TABLE 4 (continued)

$|FE|$ = absolute value of actual EPS (from I/B/E/S unsplit-adjusted actuals file) less the last mean consensus forecast prior to the earnings announcement (from I/B/E/S unsplit-adjusted summary file) scaled by the stock price at the end of quarter (Compustat quarterly item 14);
$|SPITEM|$ = absolute value of special items (Compustat quarterly item 32) divided by ending total assets (Compustat quarterly item 44);
$|DISCOP|$ = absolute value of discontinued operations (Compustat quarterly item 33) divided by ending total assets (Compustat quarterly item 44);
$M\&A^{CC}$ = number of merger and acquisition related words (acquisition, acquisitions, acquisition's, M and D, merged, merger, merging, preacquisition, synergies, synergy, and takeover) in the conference call divided by the total number of words in the conference call × 100;
$RESTRUCT^{CC}$ = number of restructuring related words (layoffs, restructure, restructured, restructuring, restructurings, severance) in the conference call divided by the total number of words in the conference call × 100;
$\Delta SALES$ = current quarter sales (Compustat quarterly item 2) less sales in the same quarter one year earlier, scaled by total assets (Compustat quarterly item 44);
AF = number of analysts who issued an EPS forecast for the current quarter (from I/B/E/S summary file); and
LNA = log of total assets (Compustat quarterly item 44).

factors. Thus, we include several control variables to capture unexpected economic events that occur during the quarter because when significant events occur, we expect managers to make expanded disclosures during the presentation and discussion sessions, leading to greater information content. The variables used are as follows:

- $|RET^{QTR}|$ = the absolute value of returns during the quarter;
- $|RET^{DAYB4}|$ = the absolute value of returns during the 24 hours prior to the call;
- $|FE|$ = the absolute value of the forecast error;
- $|SI|$ = the absolute value of special items;
- $|DISOP|$ = the absolute value of discontinued operations;
- $M\&A^{PRES}$ = the proportion of merger-and-acquisition-related words in the presentation (× 100); and
- $RESTR^{PRES}$ = the proportion of restructuring-related words in the presentation (× 100).

We include $|RET^{DAYB4}|$ in this analysis as a proxy for unexpected economic events but, as discussed previously, it is also possible that the information content of the presentation and discussion are affected by news in the press release, irrespective of any additional disclosures made during the call (i.e., because of delayed price reactions).

In addition, we include firm indicator variables to control for firm-specific differences in the informativeness of presentations and discussion periods. As discussed previously, it is possible that firms differ in the extent to which investors participate and trade during the conference call. Moreover, we are primarily interested in how manager and analyst behavior changes across time with changes in firm performance. By running our tests using firm fixed effects, we are explaining *within-firm* variation in the information content of presentations and discussion periods, an area previously unexplored in the literature.[24]

Finally, while we include firm fixed effects in our model (which controls for *static* firm-specific characteristics), doing so does not eliminate the impact of firm-specific characteristics that could possibly change over our three-year sample period.[25] Because prior research has shown some such characteristics to be related to disclosure, we include three additional control variables:

- $\Delta SALES$ = quarterly growth in sales scaled by total assets (a proxy for growth);
- AF = number of analysts who issued an EPS forecast for the current quarter (a proxy for analyst following);[26] and
- LNA = log of total assets (a proxy for firm size).

Table 4, Panel A provides univariate statistics on these variables. The mean *ROA* for the sample is less than 1 percent. The mean return over the prior quarter is −0.8 percent, and returns in the 24 hours prior to the call are, on average, close to zero (0.1 percent). However, the majority of firms met or exceeded analysts' expectations; only 28.5 percent missed expectations.

---

[24] In addition, in sensitivity analysis, we also examine length of the presentation and discussion session as an alternative proxy for information content. Controlling for firm effects is perhaps more important in this setting because firms likely vary in the way they disclose information, which likely affects the length of each segment of the call. We note, however, that in controlling for firm fixed effects rather than manager-specific effects we are presuming that disclosure policy is set by the firm and is not manager-specific. To the extent manager-specific idiosyncratic "styles" impact disclosure policies at the firm level, the power of our tests will be reduced in the event of management turnover. However, we do not believe this will bias our tests unless firms systematically hire managers with predispositions toward greater disclosure in quarters when firm performance is poor. Even if this is the case, this fact is not contrary to our hypothesis that a firm will increase disclosure when performance is poor. It simply suggests that the mechanism by which they accomplish this goal is by hiring managers with a particular disclosure "style."

[25] A firm-fixed model likely reduces the impact of these variables significantly because the cross-sectional variation in these variables is far greater than the within-firm variation.

[26] Using the quintile rank of analyst following rather than the number of analysts following results in inferentially similar results.

Table 4, Panel B presents a correlation matrix of our explanatory and control variables, after subtracting the within-firm mean from each variable (Spearman correlations reported above the diagonal and Pearson correlations reported below the diagonal). Not surprisingly, the four performance variables are correlated with each other, particularly *ROA* and *MISS* ($r = -0.18$) and *MISS* and RET^{DAYB4} ($r = -0.27$). *ROA* is also negatively correlated with $|SPITEM|$, indicating that the larger magnitude exclusions tend to be negative (or income-decreasing) in nature.

We test our hypotheses based on the following regressions:

$$\begin{aligned}|ABRET^{PRES}| = {} & \beta_0 + \beta_1 ROA + \beta_2 MISS + \beta_3 RET^{QTR} + \beta_4 RET^{DAYB4} + \beta_5 |RET^{QTR}| \\ & + \beta_6 |RET^{DAYB4}| + \beta_7 |FE| + \beta_8 |SPITEM| + \beta_9 |DISOP| + \beta_{10} M\&A^{CC} \\ & + \beta_{11} RESTR^{CC} + \beta_{12} \Delta SALES + \beta_{13} AF + \beta_{14} LNA + \beta_{15} Q2 + \beta_{16} Q3 \\ & + \beta_{17} Q4 + \beta_{18} Y04 + \beta_{19} Y05 + \varepsilon; \end{aligned} \tag{3}$$

$$\begin{aligned}|ABRET^{DISC}| = {} & \beta_0 + \beta_1 ROA + \beta_2 MISS + \beta_3 RET^{QTR} + \beta_4 RET^{DAYB4} + \beta_5 |RET^{QTR}| \\ & + \beta_6 |RET^{DAYB4}| + \beta_7 |FE| + \beta_8 |SPITEM| + \beta_9 |DISOP| + \beta_{10} M\&A^{CC} \\ & + \beta_{11} RESTR^{CC} + \beta_{12} \Delta SALES + \beta_{13} AF + \beta_{14} LNA + \beta_{15} Q2 + \beta_{16} Q3 \\ & + \beta_{17} Q4 + \beta_{18} Y04 + \beta_{19} Y05 + \beta_{20} |ABRET^{PRES}| + \varepsilon. \end{aligned} \tag{4}$$

If, consistent with the uncertainty explanation, managers provide more disclosures when firm performance is poor, then we would expect negative coefficients on *ROA*, RET^{QTR}, and RET^{DAYB4} in Equation (3) and a positive coefficient on *MISS*. If managers provide more disclosures when performance is good, consistent with the credibility argument, then we would expect the opposite signs. We also expect analysts to elicit greater disclosures during the discussion session when performance is poor, suggesting negative (positive) signs on *ROA*, RET^{QTR}, and RET^{DAYB4} (*MISS*) in Equation (4). We include fiscal quarter indicator variables (*Q2*, *Q3*, and *Q4*) to control for differences across quarters in the informativeness of calls and year indicator variables (*Y04* and *Y05*) to control for time-period-specific effects that might impact the absolute returns during conference calls. We also winsorize the top and bottom 1 percent of all variables to reduce the impact of outliers. As discussed previously, we run the above model including firm fixed effects.[27]

Because our accounting- and market-based performance variables are highly correlated, we first run our regressions separately for our accounting- and market-based performance variables. We also report results with and without control variables (except for the firm, year, and quarter indicator variables). Results for presentation returns are reported in Table 4, Panel C. Performance appears to be negatively related to the information content of the presentation, consistent with poor performance increasing uncertainty and managers attempting to alleviate that uncertainty with increased disclosures. In particular, measures of unexpected performance at the earnings announcement (RET^{DAYB4} and *MISS*) are negative and statistically significant when included separately; however, only the coefficient on RET^{DAYB4} is negative and significant when both sets of measures are included simultaneously.

---

[27] Because we have multiple observations for the same firm and market reactions to conference calls potentially have a firm-specific component, there is the potential for biased standard errors. However, according to Petersen (2009), ordinary least squares with firm fixed effects results in unbiased standard errors as long as the firm effect is static. Since we have at most three years of data for each firm (and often less), we believe any firm effect is likely to be largely static. We nevertheless test the sensitivity of our results to clustering standard errors by firm. The results are similar to those reported in the study, although t-statistics are generally lower. Specifically, those t-statistics that are significant at the 5 percent level or better continue to be; however, the coefficient on $|RET^{QTR}|$ in the regressions on the difference in returns (*DIFFRET*) is no longer significant at the 10 percent level.

Results for discussion period returns are provided in Table 4, Panel D. Discussion periods also appear to be more informative when market performance is poor. Using only our accounting-based measures of performance, the coefficient on *MISS* is positive and statistically significant. When using only the market-based measures, both the coefficients on RET^{QTR} and RET^{DAYB4} are significantly negative. Including both sets of performance measures simultaneously, the coefficient on *MISS* is no longer significant, but both market-based measures continue to be significantly negative. Overall, the results are consistent with the conjecture that, when firm performance is poor, managers fail to provide adequate disclosures during the presentation, leading to discussion sessions that uncover additional information.

Similar to our prior findings, we find that the greater the information content of the earnings release ($|RET^{DAYB4}|$), the greater the information content of both the presentation and discussion. The information content of the discussion is also greater when following more informative presentations ($|ABRET^{PRES}|$). The coefficient on $|FE|$ in the presentation regression and the coefficient on $|RET^{QTR}|$ in the discussion regression are both significantly positive, also consistent with unexpected economic events leading to greater disclosure. There also appears to be fiscal quarter differences, with second and third fiscal quarters having more informative discussions. Finally, larger firms tend to have more informative presentations and discussions, and growth firms tend to have more informative discussions.[28]

While the prior results suggest that managers provide expanded but insufficient disclosures when firm performance is poor, we conduct a more direct test of this hypothesis by running the same regression as in Equation (3), but with *DIFFRET* as the dependent variable. Recall that positive values of *DIFFRET* indicate relatively more informative presentations and negative values indicate relatively more informative discussion sessions. Thus, if discussion sessions are relatively more informative than presentations when firm performance is bad (as we predict), then we would expect positive coefficients on *ROA*, RET^{QTR}, RET^{DAYB4}, and a negative coefficient on *MISS*. As before, we winsorize the top and bottom 1 percent of the variables to reduce the impact of outliers and run our regressions including firm fixed effects.

Results are reported in Table 4, Panel E. We find some evidence consistent with our prediction. The coefficient on RET^{DAYB4} is positive and significant, indicating that when returns associated with the earnings announcement are worse, the abnormal absolute returns of the discussion session tend to be greater than the abnormal absolute returns of the presentation. The coefficient on RET^{QTR} is also positive but only marginally significant (at the 10 percent level, one-tailed). However, we do not find that accounting-based performance measures are associated with the relative informativeness of the discussion versus the presentation.

We also do not find much evidence that unexpected economic events impact the relative information content of the discussion session versus the presentation. The coefficient on $|RET^{QTR}|$ is only marginally significant. We also do not find any difference in relative information content for larger firms. Larger firms have both more informative presentations and discussions, but the difference in information content is not related to firm size. We do, however, find that second and third fiscal quarters have relatively more informative discussions.

Overall, the above results suggest that, while managers provide expanded disclosures when firm performance is poor, additional information is revealed during the discussion session with analysts in these situations. This conclusion, however, depends on whether absolute returns during the relatively narrow presentation and discussion period windows are good proxies for information content. As additional corroboration for our conclusions, we also examine the length of the presentation and discussion (in words) as an alternative proxy for information content. Specifically, we replace $|ABRET^{PRES}|$, $|ABRET^{DISC}|$, and *DIFFRET* with $LENGTH^{PRES}$, $LENGTH^{DISC}$, and *DIFFLENGTH*, respectively, in our regressions.

---

[28] Note that analyst following is not significant in either regression, unlike the results in Table 3. This is due to the fact that the analysis in Table 4 includes firm fixed effects and the variation in analyst following across time for a given firm is limited.

These results are reported in Table 4, Panel F. We report only the results of the model including all performance and control variables, although the pattern of results is similar to those reported using returns. Consistent with our results using absolute returns as the proxy for information content, we find that poor firm performance is associated with longer presentations and discussions. In addition, the difference in length between the presentation and discussion increases with firm performance. This result is consistent with our prior conclusion that, when performance is poor, discussions are relatively more informative than presentations.

In summary, analysis of both absolute returns and length in words suggest that more information is released when firm performance is poor but, relatively speaking, the discussion session has greater information content in these situations.

V. ADDITIONAL ANALYSIS: CALL CONTENT AND FIRM PERFORMANCE

We conduct a final set of analyses to provide further evidence that the documented relation between information content and performance is due to real differences in the content of information disclosed. Specifically, we examine two dimensions of the content of presentations and discussions—financial (versus nonfinancial) and forward-looking (versus backward-looking)—and correlate these content measures with firm performance.[29] The typical earnings press release contains an analysis of the prior quarter's financial performance and, therefore, provides primarily backward-looking, financial information. Thus, the greater information content of the presentation/discussion periods in periods of poor performance is likely due to increased disclosure of nonfinancial, forward-looking information. However, this prediction assumes that the analysis of last quarter's performance that is provided in the earnings release is sufficient for the market. If not, it is possible that managers will be forced to expand on this analysis, resulting in greater backward-looking, financial information, particularly during the discussion session with analysts. Survey evidence in Graham et al. (2005) is consistent with this prediction. Managers interviewed indicate that when the firm meets analysts' expectations, conference calls can be devoted to discussing future prospects. However, when firms miss expectations, the conference call shifts to discussing the reason the firm missed the consensus estimate. Given the competing possibilities and the descriptive nature of this analysis, we do not provide directional hypotheses.

Our measures of content are based on word counts of financially oriented and future-oriented words during each segment of the call. We develop a customized word list of financially oriented words based on an analysis of commonly used words in conference calls.[30] Appendix A presents the list of financially oriented words used in this analysis. Our forward-looking word list is based partially on the dictionary of future-oriented words used in LIWC (Linguistic Inquiry and Word Count). However, because the software was designed to analyze texts in a general context, the list is

---

[29] While there are obviously many dimensions one could examine, we believe these are two of the more significant ones. However, because of the way the data are computed, we are unable to cross the two attributes—that is, we cannot measure the extent of financial backward-looking information versus financial forward-looking information.

[30] Commonly used linguistic analysis tools such as Diction, LIWC, and General Inquirer do not have dictionaries for financially oriented words, thereby necessitating the creation of a custom word list. To create our list we first randomly selected 900 conference calls and identified all words occurring more than 10 times. From these 7,892 words, we identify words that we consider financial in nature.

quite short and very general (primarily "shall" and "will").[31] Thus, we augment the LIWC dictionary with other words that are often used in a business context to signify a discussion of the future (despite the fact that the words are technically present tense verbs; e.g., expects, anticipates, intends).[32] Appendix B presents the complete list of future-oriented words used in this analysis. Using these two word lists we compute the percent of financially oriented and future-oriented words in each segment of the call as a percent of the total words in each segment ($\%FIN^{PRES}$, $\%FIN^{DISC}$, $\%FUTURE^{PRES}$, $\%FUTURE^{PRES}$).[33]

To examine whether managers provide more or less financially or future-oriented disclosures when firm performance is poor, we estimate regressions similar to Equation (3), except using our *%FIN* and *%FUTURE* variables as dependent measures. We again winsorize the top and bottom 1 percent of all variables and include firm fixed effects in our analysis. If managers focus their disclosures on nonfinancial, future-oriented activities when firm performance is poor, then the coefficients on *ROA*, RET^{QTR}, and RET^{DAYB4} will be positive in the *%FIN* regressions and negative in the *%FUTURE* regressions, while the coefficient on *MISS* will be negative in the *%FIN* regressions and positive in the *%FUTURE* regressions. In contrast, if analysts actually prefer backward-looking, financially oriented disclosures when performance is poor, then we would expect opposite signs on these variables, particularly in analyses of the discussion sessions.

Table 5 reports the results of these regressions. The first set of columns present results from analyzing the percent of financially oriented words in the presentation and discussion. We find evidence that managers provide more financially oriented disclosures when firm performance is good. The coefficient on *ROA* is significantly positive, while the coefficient on *MISS* is significantly negative. Thus, better accounting performance for the quarter is associated with more financially oriented disclosures in the presentation. However, we do not find a positive association between firm performance and financially oriented disclosures during the *discussion* portion of the call. In fact, the coefficient on RET^{DAYB4} is negative and marginally significant ($p = 0.07$, two-tailed), suggesting *more* financially oriented disclosures during the discussion session when firms report negative news at the earnings announcement. This result is consistent with survey results reported in Graham et al. (2005), which indicate managers' perceptions that, when firms miss market expectations, conference call discussions focus on explaining that poor performance.

The next set of columns in Table 5 presents results of our analyses on future-oriented disclosures. We find that managers provide more future-oriented disclosures when firm performance is poor. The coefficients on *ROA* and RET^{DAYB4} are negative and highly significant (all p's < 0.001), while the coefficient on RET^{QTR} is negative and marginally significant (p <

---

[31] The LIWC software (also used in Li 2008) is available at http://www.liwc.net. We thank Professor James Pennebaker (The University of Texas at Austin), one of the developers of LIWC, for sharing the word lists for future-orientation used in LIWC. Based on discussions with Professor Pennebaker, only a handful of verbs in the English language signal future tense; thus, the relatively short list included in the LIWC dictionary.

[32] Earnings press releases often include a disclaimer about forward-looking statements that includes a list of words that management believes signify such statements. Examining a random sample of press releases, we identify a list of words used in such disclaimers and augment the list of words used by LIWC. Recent studies such as Loughran and McDonald (2009) and Henry and Leone (2009) argue that word lists created for other disciplines often misclassify and/or omit common business terms; thus, customized lists perform better in business contexts.

[33] On average, a higher percentage of words spoken during the presentation (3.9 percent) are financially oriented than during the discussion (1.7 percent), and in 97.2 percent of firm-quarters the presentation has more financially oriented words than the related discussion session. In contrast, presentation periods have somewhat lower levels of forward-looking disclosures (means of 2.1 percent and 2.4 percent for presentation and discussion, respectively). While the difference is small in magnitude (–0.004), in 72.2 percent of firm-quarters the discussion session has more future-oriented words than the presentation, which is highly significant using a Chi-square test. The fact that presentations contain a much higher percent of financially oriented words is perhaps not surprising because presentations generally contain a recap of the firm's performance during the quarter and much of this discussion is financially oriented in nature.

TABLE 5

Analysis of Presentation and Discussion Content

	Financial Orientation[a]				**Future Orientation**[b]			
	$\%FIN^{PRES}$		$\%FIN^{DISC}$		$\%FUTURE^{PRES}$		$\%FUTURE^{DISC}$	
Variable[c]	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**
ROA	**0.0157**	**3.32**	0.0028	0.79	**−0.0076**	**−2.61**	**−0.0063**	**−1.95**
RET^{QTR}	0.0007	1.50	−0.0001	−0.31	*−0.0005*	*−1.81*	0.0002	−0.77
RET^{DAYB4}	0.0010	0.64	*−0.0020*	*−1.81*	**−0.0031**	**−3.31**	−0.0007	−0.70
$MISS$	**−0.0004**	**−2.19**	0.0000	0.04	0.0010	0.97	0.0000	−0.23
$\lvert RET^{QTR}\rvert$	−0.0006	−0.89	0.0008	1.50	**0.0012**	**2.76**	−0.0003	−0.56
$\lvert RET^{DAYB4}\rvert$	−0.0018	−0.85	0.0011	0.71	**0.0026**	**1.95**	**−0.0033**	**−2.27**
$\lvert FE\rvert$	*0.0182*	*1.75*	**0.0164**	**2.11**	*−0.0106*	*−1.66*	0.0045	0.64
$\lvert SPITEM\rvert$	**0.0361**	**3.44**	*0.0139*	*1.78*	−0.0040	−0.62	−0.0011	−0.15
$\lvert DISOP\rvert$	**0.0696**	**2.21**	**0.0559**	**2.38**	−0.0049	−0.25	−0.0322	−1.50
$M\&A^{CC}$	0.0201	0.43	**0.0734**	**2.12**	0.0038	0.13	**0.1317**	**4.17**
$RESTR^{CC}$	**0.5390**	**5.62**	**0.3034**	**4.24**	**0.1697**	**2.89**	**0.2060**	**3.15**
$\Delta SALES$	0.0045	0.21	0.0021	1.30	−0.0008	−0.59	−0.0003	−0.21
AF	**−0.0001**	**−2.15**	**−0.0001**	**−4.28**	0.0000	0.37	0.0000	0.50
LNA	*−0.0008*	*−1.69*	0.0004	1.32	0.0001	0.42	−0.0005	−1.46
$Q2$	**0.0008**	**4.48**	0.0000	−0.21	**−0.0007**	**−6.31**	0.0000	0.11
$Q3$	**0.0008**	**4.33**	−0.0002	−1.34	**−0.0005**	**−4.28**	**0.0005**	**3.60**
$Q4$	**0.0019**	**10.31**	**0.0013**	**9.40**	**−0.0004**	**−3.94**	**0.0064**	**5.05**
$Y04$	−0.0001	−0.74	**−0.0065**	**−4.96**	0.0000	−0.14	0.0001	0.79
$Y05$	**−0.0007**	**−3.35**	**−0.0011**	**−7.51**	*−0.0002*	*−1.63*	−0.0001	−0.82
Firm Fixed Effects	Included		Included		Included		Included	

---

[a] $\%FIN^{PRES}$ and $\%FIN^{DISC}$ are the percent of financially oriented words in the presentation and discussion, respectively (see Appendix A for word lists).

[b] $\%FUTURE^{PRES}$ and $\%FUTURE^{DISC}$ are the percent of future-oriented words in the presentation and discussion, respectively (see Appendix B for word lists).

[c] See Table 4 for variable definitions. Bolded (*italicized*) values are significant at the 5 percent (10 percent) level using a two-tailed test.

0.10). These results are consistent with the notion that managers wish to emphasize the future in their disclosures rather than discuss the poor performance over the past quarter. We also find some evidence that discussion sessions have more future-oriented disclosures when firm performance is poor; the coefficient on *ROA* is negative and significant.

Our control variables also provide some interesting observations. It appears that unusual or unexpected events generally result in not only more financially oriented, but also more future-oriented disclosures. We also note that the first fiscal quarter has far less financially oriented disclosures during the presentation than the other three quarters (which is also true for the discussion but only for the fourth quarter). In contrast, future-oriented disclosures are *greatest* during the first fiscal quarter presentation relative to the other three quarters. This is not, however, true for discussion periods: future-oriented disclosures are actually greater in the third and fourth fiscal quarters.

In summary, we find that during the presentation managers provide less financially and more future-oriented disclosures when firm performance is poor, consistent with the idea that managers avoid discussing the poor financial performance over the prior quarter. Since earnings announcement press releases generally consist of financial information regarding the past quarter, presentations with high (low) percentages of financially (future-) oriented disclosures likely have greater overlap with the information in the earnings release.[34] Thus, this result is consistent with the greater information content of presentations when firm performance is poor.[35] For the discussion portion of the call, we find only weak evidence of a correlation between the type of disclosure and firm performance, perhaps because the extemporaneous nature of discussion periods make them more difficult to categorize.[36] Overall, the fact that we find an association between the content of conference calls and firm performance supports the notion that our prior finding of an association between the information content of calls and firm performance is due to real changes in the type of disclosures made during the presentation and the ensuing discussion.

VI. CONCLUSION

In this study, we examine the information content of the presentation and discussion portions of conference calls to provide evidence on the source of the incremental informativeness of conference calls over the accompanying press release. We find that both presentations and discussions are incrementally informative, but that discussion periods have greater information content than presentations. This finding suggests that the bigger benefit of conference calls comes from analysts' involvement in the question-and-answer session. Consistent with this conjecture, we find that the difference in information content is related to analyst following, with discussion period information content increasing in analyst following.

We also find that managers provide more information during the presentation when firm performance is poor, but the discussion period is relatively more informative than the presentation in these circumstances. This result is consistent with managers either being (1) unable to anticipate the information needs of analysts in the case of bad performance or (2) unwilling to voluntarily disclose this information in the presentation. As a result, the discussion period is relatively more informative.

Finally, we explore the types of disclosures managers make during the presentation and discussion sessions. We find that managers provide less financially oriented and more future-oriented disclosures when firm performance is poor, which is consistent with managers focusing more on nonfinancial, future-oriented disclosures when prior quarter's performance was poor. To the extent that markets find nonfinancial-oriented disclosures more informative, these findings help to explain the greater information content of presentations when firm performance is poor.

Overall, our results suggest that one of the primary benefits of hosting a conference call—as opposed to simply issuing an earnings press release—is due to the discussion session with analysts. This result is consistent with analysts playing an active role in shaping a firm's information environment. However, we recognize that the results in this study are tests of associations, and represent indirect evidence of the role analysts play in uncovering information during conference calls. Future studies might consider conducting more in-depth examinations of conference-call transcripts to provide more direct evidence of how analysts shape the information environment through their inquiries.

---

[34] Frankel et al. (2009) find that managers are less likely to provide earnings guidance during conference calls when performance deteriorates, which is inconsistent with our finding that future-oriented words increase when performance falls but *is* consistent with our finding that financially oriented words decrease when firm performance falls. Thus, it appears that the lack of earnings guidance when performance is poor is not due to a lack of future-oriented discussion but perhaps more the result of managers inability (or reluctance) to translate this discussion into financial terms.

[35] We do not find significant correlations between the proportion of financially or future-oriented words in the presentation or discussion and abnormal absolute returns during the presentation or discussion. In general, it is difficult to assess the information content of a particular type of disclosure because returns are a reflection of *all* disclosures made during the presentation or discussion and presentations contain many different types of disclosures.

[36] Discussion sessions likely do not fall into discrete "categories" of "financially oriented" or "future-oriented" but instead have combinations of different types of disclosures. Presentations, in contrast, are perhaps more easily dichotomized as being backward-/financially oriented (i.e., those that contain primarily a recap of prior quarter's performance) and those that contain more substantive nonfinancial, future-oriented information.

REFERENCES

Baginski, S., J. Hassell, and M. Kimbrough. 2004. Why do managers explain their earnings forecasts? *Journal of Accounting Research* 42 (1): 1–29.

Bamber, L. S. 1987. Unexpected earnings, firm size, and trading volume around quarterly earnings announcements. *The Accounting Review* 62 (3): 510–532.

Bowen, R., A. Davis, and D. Matsumoto. 2002. Do conference calls affect analysts' forecasts? *The Accounting Review* 77 (2): 285–316.

Bushee, B., D. Matsumoto, and G. Miller. 2003. Managerial and investor responses to disclosure regulation: The case of Reg FD and conference calls. *The Accounting Review* 79 (3): 617–643.

Dye, R. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23 (1): 123–145.

Frankel, R., M. Johnson, and D. Skinner. 1999. An empirical examination of conference calls as a voluntary disclosure medium. *Journal of Accounting Research* 37 (1): 133–150.

Frankel, R., W. Mayew, and Y. Sun. 2009. Do pennies matter? Investor relations consequences of small negative earnings surprises. *Review of Accounting Studies* 15 (1): 220–242.

Gow, I., G. Ormazabal, and D. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.

Graham, J. R., C. R. Harvey, and S. Rajgopal, 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40 (1–3): 3–73.

Henry, E., and A. J. Leone. 2009. Measuring qualitative information in capital markets research. Working paper, University of Miami.

Hirst, D. E., L. Koonce, and S. Venkataraman. 2008. Management earnings forecasts: A review and framework. *Accounting Horizons* 22 (3): 315–338.

Hollander, S., M. Pronk, and E. Roelofsen. 2010. Does silence speak? An empirical analysis of disclosure choices during conference calls. *Journal of Accounting Research* 48 (3): 531–563.

Holthausen, R., and R. Verrecchia. 1990. The effect of informedness and consensus on price and volume behavior. *The Accounting Review* 65 (1): 191–208.

Hutton, A., G. Miller, and D. Skinner. 2003. The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41 (5): 867–890.

Kimbrough, M. 2005. The effect of conference calls on analyst and market under-reaction to earnings announcements. *The Accounting Review* 80 (1): 189–219.

Lang, M., and R. Lundholm. 1993. Cross-sectional determinants of analysts' ratings of corporate disclosures. *Journal of Accounting Research* 31 (2): 246–270.

Lee, C. M. C. 1992. Earnings news and small traders; An intra-day analysis. *Journal of Accounting and Economics* 15 (2–3): 265–302.

Lee, C. M. C., M. J. Ready, and J. Seguin. 1994. Volume, volatility, and New York Stock Exchange trading halts. *Journal of Finance* 49 (1): 183–214.

Lev, B., and S. Penman. 1990. Voluntary forecast disclosure, nondisclosure, and stock prices. *Journal of Accounting Research* 28 (1): 49–76.

Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45 (2–3): 221–247.

Libby, R., J. E. Hunton, H. Tan, and N. Seybert. 2008. Relationship incentives and the optimistic/pessimistic pattern in analysts' forecasts. *Journal of Accounting Research* 46 (1): 173–198.

Loughran, T., and B. McDonald. 2009. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. Working paper, University of Notre Dame.

Mayew, W. 2008. Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research* 46 (3): 627–659.

Mayew, W., N. Sharp, and M. Venkatachalam. 2009. Are there private information benefits to participating in a public earnings conference call? Working paper, Duke University and Texas A&M University.

Mayew, W., and M. Venkatachalam. 2009. The power of voice: Managerial affective states and future firm performance. Working paper, Duke University.

Miller, G. S. 2002. Earnings performance and discretionary disclosure. *Journal of Accounting Research* 40 (1): 173–204.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.

Roychowdhury, S., and E. Sletten. 2011. Voluntary Disclosure Incentives and Earnings Informativeness. Working paper, MIT Sloan School of Management.

Skinner, D. J. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32 (1): 38–60.

Soffer, L., S. R. Thiagarajan, and B. Walther. 2000. Earnings preannouncement strategies. *Review of Accounting Studies* 5: 5–26.

Tasker, S. 1998. Bridging the information gap: Quarterly conference calls as a medium for voluntary disclosure. *Review of Accounting Studies* 3 (1–2): 137–167.

Verrecchia, R. E. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5 (December): 179–194.

APPENDIX A

List of Financially Oriented Words

accounting
accrual
accruals
accrued
allowance
allowances
amortization
amortize
amortized
asset
assets
bond
borrowed
borrowing
borrowings
budget
budgeted
budgeting
buybacks
capex
capital
capitalization
capitalize
capitalized
cash
cent
cents
convertible
cost
costs
covenants
currencies
debentures
debt
debts
deferrals
deposit
deposits
depreciation
derivative
derivatives
dividend
dividends
dollar
dollars
earnings
ebit
ebitda
eps
equities
equity
euro
euros
expenditure
expenditures
expense
expenses
finance
financed
financial
financially
financials
financing
financings
gain
gains
goodwill
hedge
hedged
hedges
hedging
impaired
impairment
impairments
income
interest
investment
investments
lease
leased
leases
leasing
lending
leverage
liabilities
liability
liquidity
loan
loans
loss
losses
margin
margins
obligations
payable
payables
payment
payments
pound
pounds
prepaid
prepayment
prepayments
pretax
profit
profitability
profits
receivable
receivables
redeemable
refinance
refinanced
refinancing
rent
rental
rentals
repurchasing
reserve
reserves
revenue
revenues
roa
roe
roi
sales
securities
securitization
security
selling
shares
swaps
tax
taxable
taxes
unamortized
unleveraged
warrants

APPENDIX B

List of Forward-Looking Words

From LIWC[a]

be
he'll
i'll
it'll
may
might
shall
she'll
they'll
tomorrow
we'll
will
won't
you'll

Additional Words[b]

expect
expects
intend
intends
anticipate
anticipates
plan
plans
believe
believes
projects
project
looking forward
going forward
look forward
go forward
looking ahead
would
should
could

---

[a] Words designated by LIWC (see http://www.liwc.net) as future-oriented.

[b] Additional words included are based on an analysis of a sample of disclaimers included in company press releases as indicating forward-looking statements.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1415–1444

American Accounting Association
DOI: 10.2308/accr-10033

Determinants of Disclosure Noncompliance and the Effect of the SEC Review: Evidence from the 2006 Mandated Compensation Disclosure Regulations

John R. Robinson
The University of Texas at Austin

Yanfeng Xue
The George Washington University

Yong Yu
The University of Texas at Austin

ABSTRACT: We investigate the economic forces that influence noncompliance with mandatory compensation disclosures and the effect of a subsequent focused enforcement action. We utilize SEC evaluations of compensation disclosures mandated by rules adopted in 2006 to examine whether noncompliance is associated with excess CEO compensation, proprietary costs, or previous media attention. We further test whether subsequent CEO compensation declines after the SEC publicly identifies noncompliance. We construct measures of defective disclosures from SEC critiques and find that disclosure defects are positively associated with excess CEO compensation and media criticism of CEO compensation during the previous year. We find no evidence supporting the contention that compensation disclosure defects are associated with proprietary costs. Furthermore, we are unable to document that the level of disclosure

---

We thank two anonymous reviewers and Harry Evans, the editor, for many helpful suggestions. We appreciate the comments by Robert Freeman, Angela Gore, Ross Jennings, Chris Jones, Steve Kachelmeier, Bill Kinney, Young K. Kwon, Susan Kulp, Clive Lennox, John McInnis, Greg Miller, Stephen Penman, Nick Seybert, Cathy Shakespeare, Erin Towery, Hal White, Sunny Yang, and the participants of the accounting colloquiums at The University of Texas at Austin, Columbia University, the University of Michigan, Nanyang Technology University, Singapore Management University, the brown bag seminar at The George Washington University, and the D.C. five-school accounting research conference. We thank Erin Towery and Tina Wang for able research assistance. We gratefully acknowledge research support provided by their respective institutions: Professor Robinson (Red McCombs School of Business and C. Aubrey Smith Professorship), Professor Yu (Red McCombs School of Business), and Professor Xue (George Washington University).

Professor Robinson worked on this paper while serving as the Academic Fellow for the Division of Corporation Finance of the Securities and Exchange Commission. The Securities and Exchange Commission, as a matter of policy, disclaims responsibility for any private publication or statement by any of its employees. Therefore, the views expressed in this paper are those of the author and do not necessarily reflect the views of the Commission or the other members of the staff of the Commission.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: July 2009
Accepted: February 2011
Published Online: April 2011

defects identified by the SEC is associated with a reduction in excess CEO compensation in the subsequent year.

Keywords: *executive compensation; mandatory disclosures; noncompliance.*

Data Availability: *Data are available from public sources identified in the paper.*

JEL Classifications: *M52; G32; G38.*

I. INTRODUCTION

The Securities and Exchange Commission (SEC) views enforcement of financial disclosure requirements as an essential element in a reliable financial accounting regime and a central part of the SEC mission (e.g., Levitt 1998; Paredes 2009). However, despite the volume of disclosure research in accounting over the past three decades, relatively little is known about why firms choose not to comply with required disclosures. Moreover, there is little direct evidence demonstrating the impact of enforcement activities by the SEC. Why firms choose not to make certain required disclosures and whether SEC enforcement is effective are important questions under a mandatory disclosure regime. In this study, we use the executive compensation regulations enacted in 2006 as a context to investigate the economic forces that influence the degree of noncompliance with mandated disclosures and the effect of subsequent SEC enforcement actions.

The 2006 executive compensation regulations provide a good setting to examine noncompliance and enforcement for multiple reasons. These regulations mandate new detailed disclosures of executive compensation. The SEC expected these new disclosures to help investors better understand and monitor firms' compensation practices (Cox 2006). Shortly after the effective date of the regulations, the SEC conducted examinations focused exclusively on executive compensation disclosures. At least two examiners were involved in each "focused study" review, and the reviews were conducted over a ten-month period between April 2007 and February 2008. This unexpected set of examinations was conducted on proxy statements issued by a large sample of public firms. The results of this focused study reveal substantial cross-sectional variation in the degree of compliance with the new regulations (White 2007), and represent a unique opportunity to (1) isolate and test for specific noncompliance determinants in an important context, and (2) examine the effectiveness of subsequent SEC enforcement actions in improving the compliance by firms whose filings are reviewed by the SEC.

Our sample contains 336 firms whose proxies were reviewed in the SEC focused study. The SEC critiques indicate that those firms' proxies contain substantial numbers of disclosure defects. Although the degree and nature of noncompliance vary substantially across firms, none of these firms appear to completely comply with the disclosure requirements. Statements from the SEC staff indicate that the staff did *not* intentionally select firms that are prone to be noncompliant for the focused study and that firms were essentially selected randomly. We independently reviewed a random sample of 50 firms not reviewed by the SEC to provide evidence on the generalizability of the disclosure defects in the SEC sample. We find very similar patterns of disclosure defects in this random sample. We also examine whether firms in the two samples corrected their disclosure defects in the subsequent year. We find that, while the firms reviewed by the SEC corrected all the identified disclosure defects, the firms not reviewed by the SEC exhibited little change in their compliance levels. This finding is consistent with the SEC's conclusion (Parratt 2009) that noncompliance with the new regulations persists until defective disclosures are publicly identified in SEC reviews. This persistence also suggests that the majority of disclosure defects are not due to inexperience with the new regulations, but rather a strategic choice made after weighing the costs and benefits of noncompliance.

Based on the SEC critiques, we construct measures of overall noncompliance and measures of three specific types of disclosure defects: pay-for-performance content (pay-performance), governance disclosures, and readability. We use these compliance measures to investigate the extent to which noncompliance with mandated disclosures is related to three particular economic constructs. We expect that managers will benefit more from withholding specific compensation information when they receive excess compensation or are protecting proprietary knowledge. Managers who are paid excess amounts of compensation have direct incentives to omit or obfuscate disclosures that reveal excess compensation. This incentive can be especially potent in 2006 because the new regulations require specific disclosures of incentive payments, methods of benchmarking, and discretionary compensation. Although shareholders want to better understand management incentives from enhanced disclosures, detailed disclosures about performance targets could inadvertently reveal important trade secrets. Hence, investors are unable to infer why managers choose to forego required disclosures. Last, we investigate whether prior negative media attention attenuates or exacerbates the level of noncompliance with the mandated compensation disclosures.

Our tests are based on regressions in which overall noncompliance is represented by the total number of defects identified in each SEC critique. We also estimate regressions using measures of specific defects related to pay-performance, governance, or readability. Excess Chief Executive Officer (CEO) compensation is calculated as the difference between actual CEO compensation and expected CEO compensation, following Core et al. (2008). Following Karuna (2007), we capture the proprietary costs with measures of three important determinants of industry competition, namely, product differentiation, market size, and entry costs. Consistent with Core et al. (2008), we identify negative media attention by collecting and evaluating all the press articles related to a firm's CEO compensation over the entire disclosure year. Negative media attention is present when a firm has at least one article explicitly criticizing CEO compensation.

After controlling for the complexity of the proxy statement, the authorship of each SEC critique, and various firm characteristics, we find that the number of overall defects in compensation disclosures is positively associated with excess CEO compensation. Furthermore, we find that the number of defects is positively related to previous negative media attention focused on CEO compensation. However, we find no evidence that the number of defects is related to any of our proxies for proprietary costs. Regarding specific defects, we find that defects in both pay-performance and governance disclosures increase with excess CEO compensation. We also find a weak positive association between pay-performance defects and negative media attention. In contrast, there is little evidence that defects in readability are associated with excess CEO compensation or negative media attention. None of the specific defects is associated with proprietary costs. In sum, our evidence indicates that the extent of noncompliance with required compensation disclosures increases with excess CEO compensation and negative media attention, but is not related to proprietary costs.

We next examine whether public disclosure of defects identified by the SEC reduces excess CEO compensation in the subsequent year. We calculate the change in total CEO compensation between fiscal years 2006 and 2007 and decompose this change into the change in expected compensation and excess compensation. We then estimate separate regressions of the change in CEO excess compensation on the overall level of noncompliance or the three specific types of disclosure defects. After controlling for mean-reversion of the excess compensation, we are unable to document that the change in excess compensation is associated with the number of total or specific types of defects. Thus, we find no evidence that the public disclosure by the SEC of defects in mandated compensation disclosures reduces excess CEO compensation.

Our results are robust to the following sensitivity analyses. First, we corroborate our inferences by using governance variables representing the level of CEO oversight as an alternative measure of excess CEO compensation. Second, we test for proprietary costs by examining the competitiveness of the CEO labor market and by identifying firms in our sample that avoided required disclosures by asserting that these disclosures would result in competitive harm. We find no evidence that disclosure defects are explained by proprietary costs. Last, the results are similar after correcting for a potential sample selection bias using a Heckman (1978) two-stage model.

Our study contributes to an understanding of the nature of disclosure noncompliance and SEC enforcement. While a large body of research has studied managers' discretionary disclosure behavior, there is limited research on compliance with mandatory disclosure requirements, with a notable exception for research on compliance with auditor change reporting (e.g., Schwartz and Soo 1996; Ettredge et al. 2001). Previous studies have examined the effect of legal regimes on accounting and reporting outcomes (Holthausen 2009; Jackson and Roe 2009), but little is known about how the SEC review process impacts financial reporting or compensation practices. This lack of research is likely due to the difficulty of obtaining a relatively objective and comprehensive measure of defective disclosures. In this study, we focus on mandated compensation disclosures and utilize disclosure defects specifically identified by the SEC's focused study to investigate both the economic determinants of the overall and specific disclosure defects and the effect of the focused SEC enforcement action.

This study also contributes to the literature on how the media influences firm decisions. Researchers argue that media attention could serve to help police corporate behavior. We find that firms are more likely to have defective compensation disclosures when CEO compensation was previously criticized in the press. Although this result must be interpreted cautiously, it suggests that managers respond to public criticism of CEO compensation by withholding mandated compensation information.

Our findings are timely and relevant to policy makers, regulators, and investors. First, our findings on the determinants of noncompliance and on the specific types of disclosures firms are likely to violate should aid regulators and investors in monitoring firms' mandatory compensation disclosure practices. Second, our finding that mandated disclosures can be compromised by management incentives is directly relevant to policy makers and regulators who seek evidence about whether agency considerations create incentives for managers to withhold compensation information. Third, our failure to detect an association between compensation defects and various measures of proprietary costs suggests that the exception for proprietary secrets might be a weak justification for avoiding detailed compensation disclosures. In particular, our evidence of an association between excess compensation and pay-for-performance defects helps justify the requirement that firms disclose specific qualitative and quantitative performance targets for individual executives.

Finally, our results from examining the relation between noncompliance and subsequent CEO compensation should be carefully considered by policy makers. Core et al. (2008) argue that one reason for their failure to find that negative media attention is associated with reductions in excess CEO pay is that firms with an entrenched CEO or weak governance are not responsive to external pressure. Consistent with this argument, our results suggest that mandatory disclosures may not be sufficient to discipline the CEO compensation. Taken together, the overall implication of the evidence is that, absent the ability to influence firm policy or the composition of the board of directors, neither negative media attention nor a focused SEC enforcement action seems to be effective in disciplining CEO compensation.

The next section presents the institutional background of our research setting and develops the hypotheses. Sections III and IV describe the sample and research method, respectively. Section V presents our results. Section VI concludes.

II. BACKGROUND AND HYPOTHESES

Background

In August 2006, the SEC adopted sweeping changes to its executive compensation disclosure rules that mandate the information companies must disclose about executive pay packages in their annual proxy statements. The new rules, which were proposed in January 2006 and generated a record number of over 20,000 comment letters, represent the first comprehensive overhaul of the executive compensation disclosure regulations since 1992. The rule changes were motivated by two important and related forces. First, previous disclosure rules provided investors inadequate and sometimes even misleading information because of the dramatic evolution in executive compensation practices.[1] Second, the option-backdating scandal and highly publicized examples of large payouts and severance packages increased demand for information about executive compensation practices (e.g., Bebchuk 2006; Cox 2006; Scannell 2006; Scannell and Lublin 2006).

The revised rules require a new "Compensation Discussion and Analysis" (CDA) section to be filed with the proxy and certified by the CEO and CFO. The new CDA section must provide an explanation and analysis of all material elements of the company's compensation goals, practices, and decisions for the CEO, CFO, three other highest-paid executive officers, and the directors. The CDA disclosures must be sufficiently precise to identify material differences in compensation policies and decisions for the five named individuals. The new rules also require companies to disclose specific quantitative or qualitative performance targets used to determine bonus payouts for executives, unless such disclosure would cause competitive harm by revealing trade secrets or confidential commercial or financial information.[2]

The new rules require a narrative description of any material factors necessary to understanding the summary compensation table and the plan-based awards table. This additional disclosure is intended to help reveal how well a company aligns its executive compensation with the performance of its executives and the company as a whole. The new rules significantly expand the disclosure regarding retirement benefits, change-in-control, and other termination compensation arrangements, eliminating surprises such as the details of the "golden goodbye" (Cox 2006, para. 39). Finally, the new regulation also requires firms to disclose the procedures and processes that the compensation committee used to set compensation policies and to determine executive and director compensation, to identify compensation consultants and to describe their role in determining executive and director compensation, and to define independent directors.

To evaluate compliance with the new rules, the SEC conducted a comprehensive focused study of compensation disclosures for approximately 350 companies from proxies issued in early 2007. During a ten-month period in 2007, specialists in the Division of Corporation Finance in the SEC evaluated each proxy statement.[3] The critiques of the sample companies highlighted disclosure defects and specified required improvements to the executive compensation disclosures. Although no tabular presentation of the results is available, two themes emerged from the review and subsequent speeches by SEC personnel.[4] First, the CDA often fails to focus on how and why a company arrives at specific executive compensation decisions and policies. Specifically, John White, Director of Division of Corporation Finance, concluded: "Far too often, meaningful analysis is missing—this is the biggest shortcoming of the first-year disclosures" (White 2007, para. 10). Second, the SEC concluded that the companies often did not use plain English or tables and graphs organized in a way to help readers understand the disclosures. The study suggested that techniques such as providing an executive summary or creating tables or charts tailored to a company's particular executive compensation program would make the disclosure more useful and meaningful (White 2007).

---

[1] Former SEC chairman Cox (2006, para. 31) notes: "This is the first time in 14 years that the Commission has undertaken significant revisions of its rules for executive compensation. Simply put, our rules are out of date, and it's high time we updated them ... Over the last decade and a half, the compensation packages awarded to directors and top executives have changed substantially. Our disclosure rules have not kept pace with changes in the marketplace. And in some cases, today's disclosure regime fuzzes up rather than illuminates the true picture of compensation."

[2] However, if the company withholds this confidential information, it must then discuss how difficult or likely it will be for executives to achieve the undisclosed performance targets. The SEC also requires the company to explain why the particular performance targets involve trade secrets or confidential commercial and financial information, and why disclosure would result in competitive harm.

[3] A minimum of two specialists reviewed each proxy. The review process is described at: http://www.sec.gov/divisions/corpfin/cffilingreview.htm.

[4] See *Staff Observations in the Review of Executive Compensation Disclosure, Division of Corporation Finance*, available at: http://www.sec.gov/divisions/corpfin/guidance/execcompdisclosure.htm.

Theory and Hypotheses

After the new regulations became effective, managers had to disclose mandated executive compensation information or risk being identified as noncompliant by the SEC and/or investors.[5] In theory, universal compliance would be achieved if disclosures could be objectively verified in a straightforward manner. This full disclosure equilibrium is similar to that characterized in Grossman's (1981) model, in which adverse selection forces firms to provide full disclosure because of costless *ex post* verification. If noncompliance could be costlessly detected, the SEC could impose sanctions against the noncompliant firms and investors could discount their stock prices. To the extent that the costs of these penalties are sufficiently high, all firms should choose to fully comply.

However, the SEC found many firms did not fully comply with the new compensation disclosure rules (Scannell and Lublin 2008; White 2007; Parratt 2009), implying that practical constraints impeded both the SEC enforcement process and market discipline. There are at least two reasons why this could have occurred with mandated compensation disclosures. First, existing SEC practices indicate that verification is costly and penalties are relatively limited.[6] While managers anticipate a periodic SEC review of all their corporate filings (at least every three years), these "normal" reviews do not typically involve specialists and would not necessarily focus on proxy disclosures and the new executive compensation disclosures in particular.[7] Hence, it is unlikely that the SEC would initially detect specific compensation disclosure defects. More importantly, managers would likely perceive that the probability of any SEC sanction for noncompliance would be low. Although the SEC must periodically review filings under the Sarbanes-Oxley Act of 2002, SEC reviewers typically only request that firms correct defective disclosures in subsequent filings.

Second, managers may decide that any valuation penalty imposed by the market on noncompliance is likely to be constrained by the inability of investors to identify and assess the reason for defective disclosures. Noncompliance could be motivated either by managerial self-interest (e.g., hide excess compensation) or to protect the firm (e.g., avoid disclosing trade secrets). Hence, investors could not ascertain whether defective disclosures signal agency problems (or weak governance) or indicate that managers are avoiding the proprietary costs of disclosure. This uncertainty about the nature of the information being withheld would prevent the market from drawing a negative conclusion from noncompliance and discounting firm value accordingly (Verrecchia 1983). Thus, if the market cannot evaluate the reasons for noncompliance and SEC enforcement is constrained, then managers who expect to benefit from withholding the required compensation information would have incentives not to comply. Next, we consider how managerial incentives to withhold mandated disclosures vary with excess executive compensation, proprietary costs, and media attention.

---

[5] We assume that firms will not disclose false information. The potential legal cost of disclosing false information would be extraordinary.

[6] Stigler (1970, 527) notes that "the extent of enforcement of laws depends upon the amount of resources devoted to the task," and Cox et al. (2003) describe the various resource constraints imposed on the SEC.

[7] At the time of preparing 2007 proxy statements, there was no indication that the SEC would conduct a focused study of these filings. Scannell and Lublin (2008) indicate that a rumor of a focused study only began to surface in August 2007, after a number of firms received initial critiques from the focused study.

Excess Compensation

Disclosure of the detailed compensation information mandated by the regulations could trigger scrutiny by shareholders, the press, or regulators, especially if these disclosures reveal excess executive compensation. We view "excess" compensation as the difference between actual compensation and expected compensation given the performance of the firm. This scrutiny could result in reduced pay, damaged personal reputation, and diminished job security. In other words, informed shareholders could act to reduce excess compensation by improving governance mechanisms or by demanding changes to executive compensation contracts. In addition, regulators or policy makers could act to reduce excess compensation by requiring additional governance mechanisms or by legislating limits on executive compensation. In sum, when the CEO receives excess compensation, there is a direct incentive either to omit disclosures that reveal the excess compensation or to obfuscate compensation disclosures, thereby concealing the information from investors, regulators, and the media. Hence, we predict the following (stated in the alternative form):

> **H1:** The level of defects in mandatory compensation disclosures will be positively associated with excess CEO compensation.

Proprietary Costs

A second potential benefit from noncompliance is avoidance of proprietary costs. Disclosure of detailed executive compensation targets could reduce firm value because of proprietary costs (Dye 2001; Verrecchia 2001). Disclosure of detailed performance targets for highly compensated executives could increase the risk of revealing proprietary information to other parties, such as competitors and trade unions. Other things equal, firms with trade secrets related to performance targets are less likely to disclose information on their compensation contracts because of the possibility of revealing trade secrets that will benefit competitors. For example, Coca Cola makes the following argument in its response to the SEC request for additional information on 2006 performance targets:

> Disclosure of the Company's internal projections and our performance against the business plans would allow competitors to understand our business priorities, areas of emphasis, investment strategies, and expectations for particular geographic regions. Competitors and industry experts may be able to deduce the investment required to deliver certain volume or profit targets. Likewise, if we disclosed these targets after-the-fact, competitors would have access to baseline information for future projected growth. As an example, if competitors discovered that we are aiming for a significant percentage increase in volume or profit in a particular geographic area, this would provide competitors with highly valuable information that would allow them to focus their competitive efforts against the Company in that area.[8]

Indeed, proprietary costs are one justification for nondisclosure that is specifically allowed by the instructions to the revised regulations.[9] Hence, we predict the following (stated in the alternative form):

---

[8] Page 5 of letter dated October 26, 2007, responding to SEC critique.

[9] Instruction 4 to Regulation S-K, Item 402(b). It is important to note that if a firm omits required disclosures based on proprietary costs, the proxy must explain how a competitor could use the omitted information in a way that would be detrimental to the firm. In our independent audit of the SEC critiques (discussed in Section III), we find that the SEC always supplied a comment when a firm claimed an exemption for the proprietary cost reason. This rules out the possibility that our failure to find an association between disclosure defects and proprietary costs is due to the SEC's comment letters ignoring defects justified on the grounds of proprietary costs.

H2: The level of defects in mandatory compensation disclosures will be positively associated with the level of proprietary costs.

Media Coverage

Besides excess compensation and proprietary costs, another factor that is likely to affect the managers' compliance decision is negative media coverage of their executive compensation. Criticism of executive pay could damage both the reputation and position of executives by informing and focusing the attention of shareholders and regulators. However, it is unclear *ex ante* whether withholding required compensation information would be more likely to alleviate or exacerbate negative media coverage. The media face the consumer demand for both information and entertainment, and as a result, while the media play an investigative role, they also tend to sensationalize news stories (Jensen 1979; Dyck et al. 2008; Miller 2006; Core et al. 2008). Hence, the impact of media coverage on management depends on whether managers believe that disclosure of more compensation information mitigates or exacerbates press criticisms of their compensation practices. Hence, we predict the following nondirectional hypothesis on the effect of negative media coverage (stated in the alternative form):

H3: The level of defects in mandatory compensation disclosures will be associated with negative media coverage prior to the mandated compensation disclosure.

We believe the effect of media coverage on firms' disclosure compliance is an important issue because if managers react to media scrutiny by improving disclosures or reducing excess compensation, then media attention could help police corporate behavior. On the other hand, if managers respond to press criticisms by reducing disclosures, then media scrutiny could have a detrimental effect. In other words, media coverage cannot change the scope of mandated compensation disclosures, but it can affect companies' level of compliance.

Prior studies have found mixed evidence on the impact of negative media coverage on executive compensation. For example, Byrd et al. (1998) find that firms with more media coverage (i.e., more attention from the press) are less likely to voluntarily disclose compensation peer groups in their proxy statements. In contrast, Laksmana (2008) finds a positive but statistically insignificant relation between prior negative media coverage and the amount of voluntarily disclosed compensation information. Core et al. (2008) find no evidence that media criticism is associated with subsequent reduced excess CEO pay or increased CEO turnover.

Subsequent CEO Compensation

Besides facilitating compliance with the revised disclosure rules, the public release of the disclosure defects identified by the SEC could also increase board accountability and lead to subsequent reductions in excess CEO compensation. This reaction can occur for at least two reasons. First, to address the disclosure defects, firms are forced to disclose performance information, which can reveal excess compensation and trigger negative reactions from shareholders and the public. Anticipating this outcome, the board and management have an incentive to reduce excess CEO compensation. Second, the public disclosure of defects in mandatory compensation disclosures is likely to draw more attention to, and generate more scrutiny of, the firm's compensation practices, putting pressure on the firm to reduce existing excess CEO compensation. In a recent study, Chen and Johnson (2008) examine the public release of SEC critiques of disclosures in financial reports. They find reduced return volatility and trading volume associated with subsequent earnings announcements and conclude that the public release of SEC reviews reduces information asymmetry. To ascertain if the public disclosure of the SEC critiques affects subsequent CEO compensation, we test the following hypothesis (stated in the alternative form):

> **H4:** Changes in subsequent excess CEO compensation will be negatively associated with disclosure defects identified in SEC critiques.

An alternative explanation for why H4 may not hold is that excess executive compensation cannot be affected merely by disclosing its existence. Core et al. (2008) argue that one explanation for their failure to find that negative media attention resulted in reduced excess compensation is that entrenched CEOs do not feel the need to reduce their compensation in reaction to public attention. Likewise, it is conceivable that defective disclosure followed by SEC criticism may not be sufficient to result in reduced excess CEO compensation.

III. SAMPLE

Sample Selection

We test compliance with the SEC's new compensation disclosure rules using the critiques issued by the SEC in 2007 regarding firms' proxy statements published in early to mid-2007. Because virtually all the firms included in this focused study have calendar year-ends, the 2007 proxy statements report executive compensation for calendar year 2006. The SEC announced that they critiqued approximately 350 companies, and after searching the EDGAR database, we are able to identify 339 of those companies.[10] Three of these firms are dropped from our analyses because of missing compensation or financial data. Our final sample contains 336 firms.

We obtain CEO compensation and director characteristics data from Standard & Poor's ExecuComp database and RiskMetrics' director's database, respectively. We collect these data directly from proxy statements when they are not available from the above-mentioned databases. The financial data are taken from Compustat and the institutional investor's holdings are obtained from the Thomson-Reuters Institutional Holdings database. Appendix A provides variable definitions.

Disclosure Defects

Defects Identified in the SEC Focused Study

The critiques of our sample firms contain 12 comments, on average, with a minimum (maximum) of one (31). We evaluate the content of each critique. Table 1 presents the top ten disclosure defects in our sample. Column (3) presents the relevant regulation and column (4) reports the frequency.

We assess whether the quality and accuracy of the critiques are consistent across the sample and across SEC staff (i.e., individuals who were responsible for drafting the critiques). We randomly select three SEC staff and 15 companies reviewed by each of these staff members and audit these critiques. We find that the critiques are accurate and that the scope and quality of the critiques are consistent across firms.

No firms in our sample were completely compliant with the new regulations. We rule out the possibility that some completely compliant firms were included in the SEC focused study but

---

[10] We also test market reactions to the comment letters (untabulated). Because we cannot identify the exact date when the comment letters were made public, we test the market reactions in the 45 calendar days from the closing letter date. This window is used because the SEC generally posts review comments on EDGAR in a batch 45 days after completion of the review (see http://www.sec.gov/divisions/corpfin/cffilingreview.htm). We find no evidence of an association between the abnormal (size-adjusted or market-adjusted) returns and the number of defects. One possible explanation is that our tests lack power. Alternatively, this finding is consistent with investors being unable to ascertain if disclosure defects are due to an agency problem or to proprietary costs.

TABLE 1
Most Common Compensation Disclosure Defects

Rank (1)	Description of the Defect (2)	SEC Regulation Violated (3)	Number (percentage) of Violations in SEC Sample of 336 Firms (4)	Number (percentage) of Violations in Random Sample of 50 Firms (5)
1	Failed to disclose the performance targets used in determining top executives' bonuses.	402(b)(2)(v)	249 (74%)	42 (84%)
2	Failed to explain how individual performance is used to determine each officer's compensation.	402(b)(2)(vii)	234 (70%)	36 (72%)
3	Failed to identify and explain material differences in compensation between top executives.	Section II.B.1 of Release No. 33-8732A	223 (66%)	39 (78%)
4	Failed to explain how the payment benefit levels are determined/negotiated for purposes of terminating top executives' employment.	402(b)(1)(v), 402(j)	181 (54%)	34 (68%)
5	Failed to explain all the functions played by the compensation consultants hired by the firm, including research, analysis, and recommendations. Also, failed to disclose the identity of the consultants.	407(e)(3)(iii)	166 (49%)	14 (28%)
6	Failed to disclose companies used as benchmarks or the composition of industry groups used as benchmarks or benchmarks used for different elements of compensation.	402(b)(2)(xiv)	143 (42%)	27 (54%)
7	Failed to describe how each element of compensation (e.g., compensation deferral) is determined and how it fits into the overall objectives and affects other elements of compensation.	402(b)(1)(v) and (vi)	119 (35%)	17 (34%)
8	Failed to explain the standard for determining material transactions and for determining whether a transaction is fair, and the procedure for review.	404(b)	106 (32%)	16 (32%)
9	Failed to explain specific factors and criteria considered in determining elements of long-term equity awards and future payout.	402(b)(1)(v)	100 (30%)	18 (36%)
10	Failed to disclose where percentile of each compensation element is targeted against benchmark companies or how tally sheet information is used.	402(b)(2)(xiv)	93 (28%)	23 (46%)

omitted from our sample. Since it is SEC policy to send a "no further comment letter" after the conclusion of any review, we conducted an EDGAR search for closing letters related to this focused study. However, the search only produced firms already included in our sample.[11]

Firms should correct their disclosure defects in the year following the SEC review for at least two reasons. First, once the defects have been publicly identified, there is little additional benefit in failing to correct the defective disclosures. Such a failure, whether intentional or inadvertent, could cast doubt on the credibility of management or organization/control of the operations. Second, a failure to correct acknowledged defects could also cause the matter to be referred to the SEC Enforcement Division for further investigation, potential penalties, and additional publicity. Also, the procedure in the Division of Corporation Finance directs personnel to review prior comment letters to confirm that defective disclosures are corrected in subsequent filings.

We conduct two additional analyses to examine the change in the compliance levels in 2007 for our sample firms. First, to verify that the SEC was satisfied with subsequent disclosures, we review subsequent SEC reviews for 30 of our sample firms selected at random. Only one of the subsequent SEC reviews contained comments on executive compensation related to the focused study. While this particular company had not completely addressed the previous comment, the SEC acknowledged that the revision met the spirit of the SEC concerns. Second, to independently verify that subsequent disclosures complied with the new regulations, we randomly select ten of our sample firms and evaluate whether the proxies issued in 2008 (for compensation paid in 2007) adequately addressed SEC comments. Our evaluation shows that each company disclosed the additional information promised in their correspondence with the SEC, with only one minor exception where the firm cited "privacy and confidentiality." Thus, the results confirm our expectation that the firms reviewed by the SEC focused study have significantly improved their subsequent compliance levels.

Defects in a Random Sample of Firms Not Reviewed by the SEC

Although the SEC has indicated that firms were not selected because of apparent noncompliance, we evaluate a random sample of firms not reviewed in the SEC focused study to provide direct evidence on the generalizability of the disclosure results from our sample.[12] We randomly select 50 firms with available proxy statements that were not reviewed by the SEC during 2006 or 2007. We evaluate each proxy statement to identify violations of the disclosure regulations after limiting this time-intensive review to the top ten defects reported in Table 1. We find that no firm in this random sample was completely compliant (i.e., all 50 firms had at least one violation). Further, we find that the pattern of the top ten defects for this random sample, presented in Table 1, column (5), is comparable to that for our sample. The mean number of top ten violations is 5.3 for this random sample versus 4.8 for our sample.

We next examine whether the firms in this random sample corrected their disclosure defects in the subsequent year. We evaluate the compensation disclosures in the proxy statements issued by these firms in 2008 (for compensation paid in 2007) for violations of the top ten disclosure defects. In contrast to the improved disclosure in our sample of 336 firms reviewed by the SEC, we find few material changes in the proxy statements for this random sample of 50 firms between 2007 and 2008 (the mean number of top ten defects for 2008 is 5.2 versus 5.3 for 2007). This result suggests that the significant improvement in the compliance levels for the firms reviewed by the SEC are due to the SEC reviews. The SEC has observed the same phenomenon in their reviews of filings made from 2006 through 2008. In her 2009 speech, Deputy Director Parratt concluded:

---

[11] In addition, we interviewed Michael Reedich (special counsel in the Division of Corporation Finance, who participated in compiling the data for the compensation study) and he stated that he was unaware of any completely compliant firms in the focused study.

[12] SEC officials indicate that firms in the 2006 focused study were not selected because of apparent noncompliance (BNA 2007). We interviewed Michael Reedich, special counsel in the Division of Corporation Finance, who participated in compiling the data for the focused study. He stated that firms were selected randomly, with additional safeguards to ensure that firms were large enough to be subject to the revised regulations and represented a broad cross-section of industries.

> The reason for ongoing comments in these areas is not because companies do not understand the disclosure requirements—to the contrary, we believe that there is a general understanding of the rules. Rather, it seems that many are reluctant to address these comment themes until we provide specific comments requesting enhanced disclosure. (Parratt 2009, para. 14)

From this evaluation, we draw three conclusions. First, the defects identified by the SEC are likely to be representative of the population of disclosure defects in executive compensation. Second, the SEC review was effective in forcing reviewed firms to correct their disclosure defects. Third, noncompliance persists after the first year of the new regulations. If the first year's noncompliance had been caused by inexperience with the new regulations, compliance would improve in the second year. Hence, the documented noncompliance apparently reflects an intentional decision not to comply rather than a temporary adjustment to the new regulations.

IV. RESEARCH DESIGN

Variable Measurement

Disclosure Defects

We construct an objective measure of overall disclosure compliance from SEC critiques. We create a variable representing the number of important defects identified by the SEC staff by recording the total number of comments in the critique (*NDEFECT*).[13] Although our independent audit of the SEC critiques revealed that the classification of defects is highly consistent for each staff member, we noted that the classification of defects differs across different staff members.[14] Hence, it is important to control for the differential styles of organizing critiques across staff members when using the number of comments (*NDEFECT*) to gauge firms' noncompliance with the disclosure rules. Thus, we include an indicator variable for each staff member to control for any staff-specific effects.

To trace the specific sources of the overall defects, we further categorize the SEC critiques into three specific types of disclosure defects: pay-performance, governance, and readability. We first identify defects in the disclosures of the content of compensation contracts from the SEC critiques and refer to these as defects in the pay-performance component of the mandated disclosures. Examples of these defects include failure to identify the benchmarks used to determine bonuses, failure to explain the reasons for material differences in compensation between executives, and failure to disclose specific components of the performance targets used to determine bonuses.[15] Second, we identify defects in the disclosures about the firm's corporate governance mechanisms related to compensation. Examples of these defects include failure to disclose the standards for identifying related-party transactions, failure to disclose the CEO's role in determining the pay for other executives, and failure to disclose how "independence" is defined and determined for individual directors. Third, we identify defects in the readability of the compensation disclosures. Examples of this type of defects include failure to limit discussion of compensation tables to information necessary to understand the tables, failure to use proper names for compensation plans to allow readers to track descriptions throughout the proxy, and failure to limit use of legalistic and confusing wording in the discussion. We count the number of comments in each of the three categories and use three variables, *PDEFECT*, *GDEFECT*, and *RDEFECT*, respectively, to represent the specific types of noncompliance.

---

[13] The results are robust to using the number of words in each critique as a proxy for the overall noncompliance.

[14] For example, one reviewer might combine defect "A" and defect "B" into one comment, while another reviewer might put these two defects into two separate comments. Because we control for the individual reviewer's impact by using reviewer fixed effects in our regressions, the personal style of each reviewer should not bias our results as long as each reviewer is consistent across her reviews.

[15] Note that the regulations do not forbid the use of discretion when granting incremental payments for performance. However, the ability to use discretion, as well as the actual exercise of discretion, must be disclosed. Item 402(b)(2)(vi) of Regulation S-K.

Excess CEO Compensation

We estimate excess CEO compensation by subtracting the expected compensation from the total compensation paid to the CEO of each sample firm. Following Core et al. (2008), we calculate the expected CEO compensation by estimating a regression of CEO compensation on the economic determinants of CEO compensation, as follows:[16]

$$\begin{aligned} Log(Compensation_i) = {} & \alpha_0 + \alpha_1 Log(Tenure_i) + \alpha_2 S\&P500_i + \alpha_3 Log(Sales_i) \\ & + \alpha_4 BM_i + \alpha_5 RET_i + \alpha_6 LagRET_i + \alpha_7 ROA_i + \alpha_8 LagROA_i + u_i, \end{aligned} \tag{1}$$

where *Compensation* is total compensation for 2006, and the remaining variables are defined in Appendix B. We then separate CEO compensation into two components: the prediction from the estimated CEO compensation regression (expected pay), and a residual (excess pay). Specifically, we compute excess compensation by estimating expected compensation for fiscal year 2006 from Equation (1) and subtracting it from actual compensation in 2006, as follows:

$$Excess\ Compensation = Compensation - Expected\ Compensation. \tag{2}$$

The method for calculating excess compensation is described in more detail in Appendix B.

Proprietary Cost Variables

Prior studies have generally used measures of industry competition as proxies for proprietary costs, based on the premise that disclosures are likely to cause more damage to a firm's competitive advantage when the firm operates in a less competitive industry and enjoys higher profitability (Harris 1998). The traditional way of measuring competition is to use a single aggregate measure of the level of concentration in the industry. Examples of such measures are the four-firm concentration ratio and the Herfindahl index (Harris 1998). Karuna (2007) draws on recent theoretical studies in economics and argues that because competition is multi-dimensional, the single industry concentration measure could be a poor proxy for competition. He proposes three measures to replace the traditional single concentration measures. We follow his approach. The three measures of the competitive environment as defined in Appendix A are *DIFF* (product differentiation), *INDSZ* (relative market size), and *ENTRY* (cost of entering the industry).[17] We expect that the competitiveness of a market or industry should increase in *INDSZ* and decrease in *DIFF* and *ENTRY*.

---

[16] Other studies that also adopt the Core et al. (2008) model in estimating expected compensation include Fischer et al. (2009), Ferri and Maber (2010), and Ertimur et al. (2011).

[17] We also conduct tests using the traditional concentration measure—the Herfindahl index—and the results (untabulated) are consistent with the results using these three measures. In addition, we test for another type of proprietary costs related to the competition in the CEO labor market (see Section V).

Media Variable

We identify negative media attention by collecting and evaluating all press articles related to a firm's CEO compensation during 2006. The articles are identified from the Factiva database following the search method described in Core et al. (2008). The indicator variable *NPRESS* represents firms that received at least one article explicitly criticizing their CEO compensation during 2006.[18] Note that our compliance measure is for proxy statements issued in early to mid-2007 regarding executive compensation paid during 2006. By measuring negative media coverage during 2006, we identify *NPRESS* prior to decisions about disclosures made in the 2007 proxy statement.

Controls

The level of disclosure compliance can be affected by various other factors. An important control is the length of compensation-related disclosures in proxy statements. Firms with more complex compensation structures should be more prone to potential disclosure problems and, thus, are more likely to have additional disclosure defects. We include the logarithm of the number of words in the proxy statements (*LPRXY*) as the control for the complexity of the compensation contracts, assuming that a more lengthy proxy statement indicates more complex compensation contracts. Byrd et al. (1998) argue that additional compensation for management is more defensible for firms with superior performance. We control for stock and financial performance by including in the regression buy-and-hold stock returns (*RET*) and return on assets (*ROA*) for 2006. Additionally, because larger firms have more resources to comply with the disclosure requirements (Schwartz and Soo 1996), we also control for firm size using the logarithm of market capitalization (*LSZ*) at the end of 2006. Finally, as previously mentioned, we control for the difference between the individual styles of SEC reviewers by including an indicator variable for each member of the SEC staff who authored critiques.

Regression Models

We test H1–H3 by estimating the following OLS model:

$$\begin{aligned} Dep.Var_i = {} & \beta_0 + \beta_1 EXPECTCOMP_i + \beta_2 EXCESSCOMP_i + \beta_3 DIFF_i + \beta_4 Log(INDSZ_i) \\ & + \beta_5 Log(ENTRY_i) + \beta_6 NPRESS_i + \beta_7 LPRXY_i + \beta_8 RET_i + \beta_9 ROA_i \\ & + \beta_{10} LSZ_i + Indicators\, for\, SEC\, reviewers + u_i, \end{aligned} \tag{3}$$

where the dependent variable is the log of the total number of defects identified in the SEC critique (i.e., *Log*(*NDEFECT*)), a measure of overall noncompliance with the SEC rules. We also use three additional dependent variables, each of which represents a specific type of disclosure deficiency: *Log*(*PDEFECT* + 1), *Log*(*GDEFECT* + 1), and *Log*(*RDEFECT* + 1). *PDEFECT* represents defects in disclosing the specific pay-performance content of the compensation contracts, *GDEFECT* represents deficiencies in the specific disclosures of the corporate governance mechanisms for determining compensation, and *RDEFECT* represents deficiencies in the readability of the compensation disclosures. We take the log to reduce the skewness in those variables, though the results are similar if we do not use a log transformation. We add 1 to *PDEFECT*, *GDEFECT*, and *RDEFECT* to avoid taking the logarithm of 0. As robustness checks, we alternatively estimate our regressions using an Ordered Probit model and a negative binomial model (Rock et al. 2001) and find similar results.

---

[18] The results are similar if we use the number of critical articles (i.e., log(number of critical articles + 1)).

We estimate this regression with our sample of 336 firms whose 2007 proxy statements were reviewed by the SEC. Because all data are drawn from the fiscal year 2006, we omit the time subscript in the variables. Our hypotheses H1–H3 predict:

H1: $\beta_2 > 0$.
H2: $\beta_3 > 0, \beta_4 < 0, \beta_5 > 0$.
H3: $\beta_6 \neq 0$.

Our last hypothesis (H4) examines whether firms reduce excess CEO compensation after public disclosure of the SEC critiques. To test H4, we estimate the following OLS model:

$$\begin{aligned} \Delta EXCESSCOMP_i &= \varsigma_0 + \varsigma_1 Log(NDEFECT_i) + u_i, \text{ or} \\ \Delta EXCESSCOMP_i &= \varsigma_0 + \varsigma_2 Log(PDEFECT_i + 1) + \varsigma_3 Log(GDEFECT_i + 1) \\ &\quad + \varsigma_4 Log(RDEFECT_i + 1) + u_i. \end{aligned} \tag{4}$$

To calculate the dependent variable, we begin with the change in the CEO's total compensation from 2006 (pre-SEC review) to 2007 (post-SEC review). We then decompose the total change in the CEO's compensation into its two components, *ΔEXPECTCOMP* and *ΔEXCESSCOMP*. Our focus is on the change in CEOs' excess compensation because if firms that receive more criticism from the SEC reduce CEO excess compensation, we expect negative signs on the four measures of defects, $\varsigma_1 < 0$, $\varsigma_2 < 0$, $\varsigma_3 < 0$, and $\varsigma_4 < 0$.

V. RESULTS

Summary Statistics and Correlation Analysis

Table 2 presents the summary statistics of our sample. The variation among sample firms in terms of the number of defects detected by the SEC reviewers is sufficiently large for the purpose of our study. An SEC critique, on average, identifies 12 defects with a standard deviation close to four.

TABLE 2

Descriptive Statistics

Variable	Mean	Min.	25%	Median	75%	Max.	S.D.
NDEFECT	11.80	1.00	9.00	12.00	14.00	31.00	3.94
PDEFECT	8.17	0.00	6.00	8.00	10.00	22.00	2.93
GDEFECT	2.62	0.00	1.00	2.00	4.00	10.00	1.72
RDEFECT	1.00	0.00	0.00	1.00	1.00	7.00	1.18
EXPECTCOMP	8.74	5.96	8.19	8.77	9.30	10.54	0.77
EXCESSCOMP	0.00	−3.37	−0.30	0.00	0.34	1.98	0.64
DIFF	1.31	0.97	1.12	1.18	1.32	2.91	0.37
Log(INDSZ)	11.18	7.97	9.99	10.92	12.38	14.62	1.64
Log(ENTRY)	7.86	4.08	6.70	8.28	9.08	11.19	1.77
NPRESS	0.11	0.00	0.00	0.00	0.00	1.00	0.31
LPRXY	10.16	8.72	9.90	10.18	10.46	11.35	0.46
RET	0.13	−0.51	0.00	0.13	0.26	1.02	0.26
ROA	0.05	−0.31	0.02	0.04	0.09	0.27	0.07
LSZ	8.98	6.54	7.91	8.61	10.02	12.38	1.36

This table reports the descriptive statistics. The sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC.
Variable definitions are provided in Appendix A (all variables are measured as of 2006).

An average SEC critique is composed of eight comments on pay-performance disclosures, three comments on corporate governance disclosures, and one comment on readability.

Table 3 reports correlations between key variables. Consistent with H1, the total number of defects (*NDEFECT*) and the number of defects related to pay-performance disclosures (*PDEFECT*) are positively correlated with CEO excess compensation (*EXCESSCOMP*). In contrast, there is little evidence that the number of defects related to governance disclosures (*GDEFECT*) or readability (*RDEFECT*) is correlated with CEO excess compensation (*EXCESSCOMP*). The correlation results provide little support for H2 because the measures of proprietary costs (*DIFF*, *Log*(*INDSZ*), and *Log*(*ENTRY*)) are generally not correlated with any of the four measures of disclosure defects. Consistent with H3, we find that negative media coverage of CEO compensation (*NPRESS*) is positively correlated with the total number of defects (*NDEFECT*), as well as the number of defects related to pay-performance disclosures (*PDEFECT*) and governance disclosures (*GDEFECT*).

Regression Results

Table 4 presents the results from estimating Equation (3), where the dependent variable is the summary measure of firm's overall disclosure noncompliance (*Log*(*NDEFECT*)). Consistent with H1, firms with higher excess CEO compensation tend to have more disclosure defects. Inconsistent with H2 and with claims made by firms, the results provide no evidence that proprietary costs are associated with disclosure defects. The estimated coefficients on the three proprietary cost measures are all statistically insignificant. Consistent with H3, firms receiving negative media coverage regarding CEO compensation tend to have more defects in their compensation disclosures.

The impact of excess compensation and negative media coverage on disclosure defects appears to be economically significant. Based on the results in Table 4, we estimate that when the estimated excess compensation (log of actual compensation minus log of expected compensation) increases by one standard deviation (0.64), there should be 9.9 percent more defects in the firm's proxy statement. Firms with negative media coverage in the prior year have 15 percent more defects in their proxy statements than firms without negative media coverage.

Table 5 presents the results from estimating Equation (3) using the three measures gauging specific types of the disclosure defects (i.e., $Log(PDEFECT + 1)$, $Log(GDEFECT + 1)$, and $Log(RDEFECT + 1)$). The results show that excess compensation is significantly associated with defects related to two aspects of the proxy disclosures, pay-performance and corporate governance. The third type of defect, the readability of the disclosures, is not associated with excess compensation. There is also some weak evidence that pay-performance defects increase with negative media coverage.

The final hypothesis (H4) requires testing the relation between the disclosure defects and the change in CEO excess compensation in the year subsequent to the SEC focused study. Table 6, Panel A, reports the results from estimating Equation (4). Although the estimated coefficients on $Log(GDEFECT + 1)$ and $Log(RDEFECT + 1)$ are not significant, the estimated coefficient on $Log(PDEFECT + 1)$ is significant in the predicted direction. Hence, there is some preliminary evidence that the number of defects related to pay-performance is associated with a reduction in subsequent excess CEO compensation (H4).[19] We also estimate regressions (untabulated) after substituting the change in total compensation and the change in expected compensation for the change in excess compensation. We find no association between the number of defects identified in

---

[19] It should be noted that SEC critiques do not address the magnitude of executive compensation, but rather whether the required compensation information is completely and clearly disclosed.

TABLE 3

Correlation Matrix

	NDEFECT	*PDEFECT*	*GDEFECT*	*RDEFECT*	*EXPECTCOMP*	*EXCESSCOMP*	*DIFF*	*Log(INDSZ)*	*Log(ENTRY)*	*NPRESS*
NDEFECT		0.85***	0.65***	0.27***	−0.01	0.25***	0.04	0.01	0.01	0.21***
PDEFECT	0.81***		0.27***	−0.05	−0.01	0.27***	0.02	−0.03	−0.01	0.18***
GDEFECT	0.60***	0.22***		0.03	−0.05	0.10*	0.03	0.03	0.00	0.11**
RDEFECT	0.31***	−0.02	0.06		0.06	−0.00	0.03	0.06	0.05	0.10*
EXPECTCOMP	−0.01	−0.02	−0.04	0.02		0.00	−0.10	0.29***	0.34***	0.13*
EXCESSCOMP	0.18***	0.21***	0.07	0.10*	0.01		−0.00	0.02	0.02	0.19***
DIFF	0.06	0.03	0.05	−0.02	−0.02	0.04		0.06	−0.30***	0.01
Log(INDSZ)	−0.02	−0.05	−0.11**	0.01	0.30***	0.04	0.24***		0.69***	0.08
Log(ENTRY)	0.03	0.03	−0.02	0.03	0.31***	0.03	0.13**	0.76***		0.09*
NPRESS	0.22***	0.14**	0.12**	0.07	0.12*	0.21***	0.04	0.10*	0.10*	

*, **, *** Denote significance levels of 10 percent, 5 percent, and 1 percent, respectively, two-tailed.

This table reports Pearson (top diagonal) and Spearman (bottom diagonal) correlations for the regression variables. The sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC.

CALCUTTA UNIVERSITY CENTRAL LIBRARY

TABLE 4

Regression Results: Determinants of Overall Compensation Disclosure Defects

Independent Variables	**Prediction**	**Dependent Variable: *Log(NDEFECT)*** **Coefficient Estimate**	**Standard Error**
Test of Excess Compensation			
EXPECTCOMP	?	0.069	(0.043)
EXCESSCOMP	+	0.147***	(0.030)
Tests of Proprietary Costs			
DIFF	+	0.022	(0.065)
Log(INDSZ)	–	0.003	(0.019)
Log(ENTRY)	+	–0.005	(0.018)
Test of Media Coverage			
NPRESS	?	0.137**	(0.064)
Controls			
LPRXY		–0.027	(0.043)
RET		–0.003	(0.079)
ROA		0.320	(0.278)
LSZ		–0.040*	(0.024)
Reviewer fixed effects		Included	
R^2		0.30	

---

*, **, *** Denote significance levels of 10 percent, 5 percent, and 1 percent, respectively; one-tailed for coefficients with predictions, and two-tailed otherwise.

This table reports the results of testing the determinants of defects in compliance with the compensation disclosure requirements using Equation (3). The sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC. Reviewer fixed effects are indicator variables equal to 1 for each individual SEC staff member that reviewed the proxy disclosures. See Appendix A for other variable definitions.

the SEC critiques (total defects or three specific types of defects) and the change in total or expected CEO compensation.

A significant issue with the Equation (4) regression is the potential for mean reversion in CEO compensation (Core et al. 2008). For example, if a CEO has high compensation in one year, the compensation in subsequent years will tend to be lower. Because excess compensation is larger for firms with more disclosure defects, CEOs with more defective disclosures may be more likely to experience mean reversion in pay. That is, this mean reversion could introduce a negative relation between our defect measures and the change in excess compensation. In Panel B of Table 6, we control for mean reversion by including the 2006 estimate of excess compensation (*EXCESSCOMP* 2006) in Equation (4). The results from estimating this alternative specification do not support an association between the level of disclosure defects and changes in subsequent excess CEO compensation.

Unfortunately, we cannot be certain if the results in Panel A of Table 6 are due to mean reversion or whether mean reversion compromises the power of our test to detect the effect of the defective disclosures in Panel B of Table 6. In other words, while the reduction in excess CEO compensation after release of the SEC compensation critiques suggests that SEC critiques could play a disciplinary role, we cannot claim to have documented such an association because mean reversion in CEO compensation is an alternative explanation. We also note that our failure to support H4 is consistent with the evidence in Core et al. (2008) that negative media attention does not affect CEO compensation.

TABLE 5

Regression Results: Determinants of Specific Compensation Disclosure Defects

Independent Variables	Prediction	Dependent Variables: Log (*PDEFECT* + 1)	Dependent Variables: Log (*GDEFECT* + 1)	Dependent Variables: Log (*RDEFECT* + 1)
Test of Excess Compensation				
EXPECTCOMP	?	0.068	0.020	−0.031
		(0.040)	(0.064)	(0.068)
EXCESSCOMP	+	0.148***	0.092**	0.020
		(0.028)	(0.045)	(0.047)
Test of Proprietary Costs				
DIFF	+	0.035	0.015	0.047
		(0.061)	(0.097)	(0.103)
Log(*INDSZ*)	−	−0.007	0.004	0.007
		(0.018)	(0.029)	(0.031)
Log(*ENTRY*)	+	−0.003	−0.001	0.008
		(0.017)	(0.027)	(0.029)
Test of Media Coverage				
NPRESS	?	0.088	0.122	0.083
		(0.060)	(0.095)	(0.101)
Control variables		Included	Included	Included
Reviewer effects		Included	Included	Included
R^2		0.30	0.15	0.09

, * Denote significance levels of 5 percent and 1 percent, respectively; one-tailed for coefficients with predictions, and two-tailed otherwise.

This table reports the results of testing the determinants of specific disclosure defects using Equation (3). The sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC. Fixed effects for SEC reviewers are indicator variables equal to 1 for each individual SEC staff that reviewed the compensation disclosures. See Appendix A for other variable definitions. The control variables (untabulated) include *LPRXY*, *RET*, *ROA*, and *LSZ*. The standard errors are reported in parentheses.

Supplemental Tests

Oversight of the CEO as an Alternative Measure of Excess CEO Compensation

Our analyses of excess CEO compensation depend upon estimates generated by the Core et al. (2008) model. Although we do not test the validity of this model, we have no *a priori* reason to suspect that the estimates of excess compensation generated from the model would be spuriously correlated with defects in compensation disclosures. To the extent that the model adds noise into the excess compensation measure, it is likely to bias against our findings. However, to provide evidence on the robustness of our results, we replicate our regressions after replacing excess CEO compensation with a series of governance variables capturing the level of oversight on CEO compensation. Core et al. (1999) show that weak corporate governance can lead to suboptimal compensation contracts for management and excess management pay.

An extremely independent board would have a non-CEO chair, more independent directors, fewer outside directors appointed by the CEO, fewer busy outside directors, and fewer gray outside directors. A relatively small board would also likely be more effective in monitoring managers (Jensen 1993). Finally, better corporate governance could also be achieved by effective external monitoring by powerful institutional investors. Following Core et al. (1999) and Hartzell and Starks (2003), we construct a measure of external monitoring, *INST5*, and six measures of the level of board independence: *CHAIR*, *INDIR*, *DIRCEO*, *DIRBUSY*, *DIRGRAY*, and *LBOARD* (defined in

TABLE 6

Regression Results: Change in Subsequent Excess CEO Compensation after SEC Critique of Compensation Disclosures

Panel A: Test of Defects

Independent Variables	Prediction	Dependent Variable: $\Delta EXCESSCOMP$	
Intercept		0.182	0.317
		(0.220)	(0.215)
$Log(NDEFECT)$	–	–0.076	
		(0.090)	
$Log(PDEFECT + 1)$	–		–0.205**
			(0.097)
$Log(GDEFECT + 1)$	–		0.085
			(0.067)
$Log(RDEFECT + 1)$	–		0.049
			(0.065)
R^2		0.00	0.02

Panel B: Control for Mean Reversion in CEO Compensation

Independent Variables	Prediction	Dependent Variable: $\Delta EXCESSCOMP$	
Intercept		–0.286	–0.182
		(0.215)	(0.214)
$Log(NDEFECT)$	–	0.119	
		(0.089)	
$Log(PDEFECT + 1)$	–		0.013
			(0.097)
$Log(GDEFECT + 1)$	–		0.098
			(0.063)
$Log(RDEFECT + 1)$	–		0.073
			(0.061)
$EXCESSCOMP$ 2006	?	–0.369***	–0.361***
		(0.052)	(0.053)
R^2		0.13	0.14

, * Denote significance levels of 5 percent and 1 percent, respectively; one-tailed for coefficients with predictions, and two-tailed otherwise.

This table reports the results of testing the impact of defects in compensation disclosures identified by the SEC review on the change in excess CEO compensation using Equation (4). The sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC. The change in the logarithm of the CEO's total compensation from 2006 to 2007 is decomposed into the change in the CEO's expected compensation from 2006 to 2007 and the change in the CEO's excess compensation ($\Delta EXCESSCOMP$) using the model described in Appendix B. See Appendix A for other variable definitions. The standard errors are reported in parentheses.

Appendix A). Thus, our supplemental tests of H1–H3 estimate the following regression model:

$$\begin{aligned} Dep.Var_i = {} & \gamma_0 + \gamma_1 CHAIR_i + \gamma_2 INDIR_i + \gamma_3 DIRCEO_i + \gamma_4 DIRGRAY_i + \gamma_5 DIRBUSY_i \\ & + \gamma_6 LBOARD_i + \gamma_7 INST5_i + \gamma_8 DIFF_i + \gamma_9 Log(INDSZ_i) + \gamma_{10} Log(ENTRY_i) \\ & + \gamma_{11} NPRESS_i + \gamma_{12} LPRXY_i + \gamma_{13} RET_i + \gamma_{14} ROA_i + \gamma_{15} LSZ_i \\ & + Indicators\ for\ SEC\ reviewers + u_i. \end{aligned} \tag{5}$$

As in Equation (3), the dependent variable takes four forms: the summary measure $Log(NDEFECT)$

and the three aspects of the disclosure defects: *Log*(*PDEFECT* + 1), *Log*(*GDEFECT* + 1), and *Log*(*RDEFECT* + 1). We expect firms with a weaker board and less external monitoring (as measured by institutional ownership) to have a higher level of noncompliance. The predicted signs on the coefficients are as follows:

H1: $\gamma_1 > 0, \gamma_2 < 0, \gamma_3 > 0, \gamma_4 > 0, \gamma_5 > 0, \gamma_6 > 0, \gamma_7 < 0$.

Table 7 presents results of estimating Equation (5) using the summary measure of firms' compliance (*Log*(*NDEFECT*)), and Table 8 presents results using the three specific aspects of the defects as dependent variables (*Log*(*PDEFECT* + 1), *Log*(*GDEFECT* + 1), and *Log*(*RDEFECT* + 1)). Both equations use corporate governance variables, rather than excess CEO compensation, as a measure of CEO overcompensation. Consistent with the results in Tables 3 and 4 (using excess compensation), we find some support for H1 and H3, but no support for H2. In Table 7, two corporate governance mechanisms stand out as significant in the predicted directions: *CHAIR* and *INST5*. Overall disclosure defects are higher when the CEO is also the chair of the board, and there is less influence from powerful outside institutional investors.

The significant coefficients on *CHAIR* and *INST5* persist in Table 8, where specific disclosure defects are examined. A firm with its CEO as the chair of the board and without enough external monitoring from powerful institutional investors is more likely to show deficiencies in disclosing information about the pay-performance content of its executive compensation contracts and corporate governance mechanisms. Further, firms with more outside directors who are busy (*DIRBUSY*) have more defects in governance disclosures, and firms with more independent directors (*INDIR*) or fewer outside directors who are appointed by the CEO (*DIRCEO*) or busy (*DIRBUSY*) have higher readability for their compensation disclosures.

CEO Labor Market as a Proprietary Cost

We examine another type of proprietary cost that a firm could incur by disclosing specific CEO performance information. If the competition for talented CEOs is high, then a firm can have an incentive not to disclose specific details of its CEO compensation contract because of the risk that other firms will use this information to appropriate its CEO. Cremers and Grinstein (2010) and Murphy and Zabojnik (2007) suggest that the competition for CEOs depends on the components of CEOs' skills and ability that are transferable across firms (non-firm-specific) versus skills that are only valuable within the firm (firm-specific).

Following this stream of prior research, we use the percentage of new CEOs that were hired from outside industries based upon Panel C of Table 4 in Cremers and Grinstein (2010) to measure the competition in the CEO labor market. We then replicate our regressions after including this new variable or substituting it for the three existing proprietary cost variables. The results (untabulated) for H1 and H3 are similar to those presented in the tables. Consistent with the reported results for H2, we find no evidence that firms who face higher competition in the CEO labor market have more defects in their mandated compensation disclosures.[20]

Refusal to Disclose Performance Targets as an Indicator of Proprietary Costs

In another set of tests for proprietary costs, we identify firms in our sample that refused to disclose specific performance goals and compensation information by arguing that these disclosures would result in competitive harm. These firms are identified from the acknowledgment letters sent by the SEC upon closing the SEC review. For these firms, the acknowledgement letters state that the SEC has no

---

[20] We also replicate our regressions after adding industry fixed effects to control for any industry-specific proprietary cost factors (including the CEO labor market). Our inferences remain unchanged.

TABLE 7

Regression Results: Determinants of Overall Compensation Disclosure Defects Substituting CEO Oversight Variables for Excess CEO Compensation

Independent Variables	Prediction	Dependent Variable: *Log(NDEFECT)* Coefficient Estimate	Standard Error
Test of CEO Oversight			
CHAIR	+	0.090**	(0.043)
INDIR	−	0.045	(0.092)
DIRCEO	+	−0.015	(0.073)
DIRGRAY	+	−0.135	(0.254)
DIRBUSY	+	0.024	(0.123)
LBOARD	+	0.004	(0.040)
INST5	−	−0.539***	(0.202)
Tests of Proprietary Costs			
DIFF	+	0.029	(0.067)
Log(INDSZ)	−	−0.001	(0.021)
Log(ENTRY)	+	−0.005	(0.019)
Test of Media Coverage			
NPRESS	?	0.191***	(0.065)
Control variables		Included	
Reviewer fixed effects		Included	
R^2		0.27	

, * Denote significance levels of 5 percent and 1 percent, respectively; one-tailed for coefficients with predictions, and two-tailed otherwise.

This table reports the results of testing the determinants of overall compensation defects using Equation (5). The sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC. Reviewer fixed effects are indicator variables equal to 1 for each individual SEC staff. See Appendix A for other variable definitions. The standard errors are reported in parentheses.

factual basis to judge whether the information those firms refused to disclose is proprietary or not. We test whether proprietary costs are higher for these firms than for others in our sample using both univariate comparisons and a logistic multivariate model. The dependent variable in the logistic model is an indicator variable for firms that refused to disclose requested compensation information on the basis of proprietary costs, and the independent variables in the model are the same as those in Equation (3). The results of the univariate comparisons and the logistic regression (untabulated) do not provide any evidence that proprietary costs are higher for the firms claiming proprietary costs.

Heckman Test and Additional Robustness Results

In this subsection, we examine whether our results are confounded by potential sample selection bias because our sample was selected by the SEC. We construct another sample of 639 firms (the *non-SEC* sample) that were not included in the SEC focused study and have the necessary compensation, governance, accounting, and price data for the fiscal year 2006.[21] We calculate the

---

[21] Our sample firms are significantly larger than the non-SEC firms, and consistent with this size differential, the CEOs of our sample firms have significantly higher expected compensation. Interestingly, the two proxies for the proprietary costs, *DIFF* and *Log(ENTRY)*, are also significantly larger for our sample firms, suggesting that the firms in our sample face higher proprietary costs than the non-SEC sample, on average. The rest of the firm characteristics, however, do not differ between the two samples.

TABLE 8

Regression Results: Determinants of Specific Compensation Disclosure Defects Substituting CEO Oversight Variables for Excess CEO Compensation

Independent Variables	Prediction	Dependent Variables: *Log* (*PDEFECT* + 1)	Dependent Variables: *Log* (*GDEFECT* + 1)	Dependent Variables: *Log* (*RDEFECT* + 1)
Test of Excess Compensation Using Governance/Oversight Proxies				
CHAIR	+	0.096***	0.135**	−0.075
		(0.041)	(0.063)	(0.066)
INDIR	−	0.110	0.022	−0.241**
		(0.086)	(0.133)	(0.140)
DIRCEO	+	−0.067	−0.044	0.166*
		(0.068)	(0.105)	(0.111)
DIRGRAY	+	−0.137	0.179	0.130
		(0.239)	(0.369)	(0.388)
DIRBUSY	+	−0.119	0.286*	0.267*
		(0.116)	(0.178)	(0.188)
LBOARD	+	−0.012	0.062	0.030
		(0.038)	(0.058)	(0.061)
INST5	−	−0.375**	−0.480*	−0.214
		(0.190)	(0.293)	(0.308)
Test of Proprietary Costs				
DIFF	+	0.039	0.018	0.051
		(0.063)	(0.097)	(0.102)
Log(INDSZ)	−	−0.011	0.006	0.012
		(0.019)	(0.030)	(0.031)
Log(ENTRY)	+	−0.001	−0.008	0.001
		(0.018)	(0.028)	(0.029)
Test of Media Coverage				
NPRESS	?	0.156***	0.134*	0.058
		(0.061)	(0.094)	(0.099)
Control variables		Included	Included	Included
Reviewer effects		Included	Included	Included
R^2		0.27	0.17	0.12

*, **, *** Denote significance levels of 10 percent, 5 percent, and 1 percent, respectively; one-tailed for coefficients with predictions, and two-tailed otherwise.

This table reports the results of testing the determinants of specific compensation defects using Equation (5). The sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC. Reviewer fixed effects are indicator variables equal to 1 for each individual SEC staff. See Appendix A for other variable definitions. The control variables (untabulated) include *LPRXY*, *RET*, *ROA*, and *LSZ*. The standard errors are reported in parentheses.

expected and excess compensation by combining our sample firms and non-SEC sample firms, and using the combined group of firms, we estimate a Heckman (1978) two-stage model to directly control for potential selection bias.

In the first stage, we model the determinants of the SEC's selection choice. We include all the independent variables in the second-stage main models except for *LPRXY*. Importantly, we define an instrumental variable that equals 1 for firms with calendar (December 31) year-ends. Because the SEC focused study started in April 2007, almost all firms included in the study have December fiscal year-ends. Specifically, only two firms in our sample have non-December year-ends, and in

TABLE 9

Tests for Generalizability

Panel A: Heckman Regression Results for Equation (3)

Independent Variables	**Prediction**	**Dependent Variable: *Log(NDEFECT)***
Test of Excess Compensation		
EXPECTCOMP	?	0.054
		(0.048)
EXCESSCOMP	+	0.143***
		(0.028)
Tests of Proprietary Costs		
DIFF	+	0.010
		(0.074)
Log(INDSZ)	–	0.007
		(0.020)
Log(ENTRY)	+	−0.008
		(0.018)
Test of Media Coverage		
NPRESS	?	0.157**
		(0.078)
Controls		
LPRXY		−0.028
		(0.042)
RET		0.008
		(0.084)
ROA		0.395
		(0.314)
LSZ		−0.046
		(0.045)
Inverse Mills Ratio		−0.057
		(0.148)
Reviewer fixed effects		Included

Panel B: Heckman Regression Results for Equation (4)

Independent Variables	**Prediction**	**Dependent Variable: Δ*EXCESSCOMP***			
Intercept		0.331	0.496**	−0.303	−0.187
		(0.242)	(0.236)	(0.229)	(0.227)
Log(NDEFECT)	–	−0.088		0.150	
		(0.096)		(0.091)	
Log(PDEFECT + 1)	–		−0.249***		0.035
			(0.104)		(0.100)
Log(GDEFECT + 1)	–		0.105		0.110
			(0.071)		(0.065)
Log(RDEFECT + 1)	–		0.057		0.082
			(0.068)		(0.062)
EXCESSCOMP 2006	?			−0.448***	−0.439***
				(0.051)	(0.051)
Inverse Mills Ratio		−0.152	−0.143	−0.086	−0.084
		(0.093)	(0.092)	(0.084)	(0.084)

(continued on next page)

TABLE 9 (continued)

, * Denote significance levels of 5 percent and 1 percent, respectively; one-tailed for coefficients with predictions, and two-tailed otherwise.

This table reports the results of testing the generalizability of our findings. The test sample consists of 336 firms whose compensation disclosures in the proxy statements for 2006 were reviewed by the SEC. The non-SEC sample includes 639 firms not critiqued by the SEC with the necessary compensation, governance, accounting, and price data for the fiscal year 2006. The expectation and excess compensations components are estimated using all the observations (i.e., both the test and non-SEC samples). Panels A and B report the results of estimating Equations (3) and (4), respectively, using the Heckman method to control for the potential selection biases arising from the SEC's selection decisions. Inverse mills ratio is estimated from the first-stage model, in which the dependent variable is an indicator equal to 1 for the test sample and 0 for the non-SEC sample, and the independent variables include all the independent variables in Equation (4) except for *LPRXY* and an indicator variable equal to 1 for firms with December year-end. See Appendix A for other variable definitions. The standard errors are reported in parentheses.

comparison, 100 (15 percent) of the non-SEC sample have non-December year-ends. Using this first-stage selection model, we generate the inverse Mills ratio and include it in our second-stage estimation of our main regression models.

The results of using the Heckman two-stage procedure to estimate Equation (3) are presented in Table 9, Panel A. The inverse Mills ratio included in the second-stage estimation is not statistically significant at conventional confidence levels. More importantly, all of our inferences remain unchanged. The coefficients on *EXCESSCOMP* and *NPRESS* remain significantly positive, and the coefficients on the proprietary cost variables are insignificant.[22] Table 9, Panel B, presents the results of using the Heckman two-stage procedure to estimate Equation (4), and column (6) of Panel B presents the Heckman two-stage results after including *EXCESSCOMP* 2006 to control for mean reversion in CEO compensation. Again, the evidence from the Heckman regressions is consistent with the results presented in earlier tables.

We conduct several sets of additional tests (untabulated) to ascertain if our test results are due to potential confounding factors. These tests include substituting various alternative estimates of excess compensation into Equation (3), as well as re-estimating the Heckman two-stage regressions.[23] The estimated results for Equation (3) are consistent with those presented in the tables, although the p-values are sometimes weaker (less significant). We also re-estimate the Heckman two-stage analyses after substituting for *EXCESSCOMP* ten indicator variables representing each decile of *EXCESSCOMP*. The results (untabulated) are almost identical to those presented in Table 9. In a final set of tests, we re-estimate Equation (3) after including an additional variable measuring total media coverage. The estimated coefficient on total media attention is statistically insignificant, but otherwise the regression results are similar to those presented in the tables. After concluding this battery of tests, we are unable to detect any confounding factor responsible for the pattern of results presented in the tables.

VI. CONCLUSION

We examine the new executive compensation disclosures mandated in 2006 as a context for investigating noncompliance with the mandated disclosures. Using SEC critiques, we measure the extent and nature of noncompliance with newly mandated compensation disclosures for a sample of 336 companies. We find that the number of overall defects, as well as defects in pay-for-performance information, is positively associated with excess CEO compensation and negative media attention. However, we find no evidence that the level of noncompliance is related to any of our measures for potential proprietary costs, and we do not observe a significant association between the change in CEO excess compensation and the number of disclosure defects. We conduct a battery of tests to substantiate that the results using the firms in the SEC focused study can be generalized to a broad population.

---

[22] We also estimate the Heckman model for the governance regression (i.e., Equation (5)) and include governance variables (rather than excess compensation) in the first-stage model. Alternatively, we include both the compensation and governance variables in the first-stage model for the excess compensation and governance regressions. In both cases, the results (untabulated) are similar to the tabulated results.

[23] These substitutes include estimates of excess compensation using various alternative measures of firm size, controls for future firm performance, estimates based on all available observations (i.e., combining the non-SEC sample with our sample), and using compensation for the top group of executives as a substitute for CEO compensation.

This study expands our knowledge about why firms fail to comply with mandated compensation disclosures. The evidence that management incentives compromise mandated disclosures until reinforced by regulators is relevant for both regulators and the legislators who formulate policy. Our results are inconsistent with the contention that negative media attention will reinforce mandated disclosures and the trade secret justification commonly offered for avoiding compensation disclosures. Finally, the failure of media attention and mandated disclosure to affect subsequent excess CEO compensation suggests that, absent the ability to influence firm policy or the composition of the board of directors, neither negative media attention nor a focused SEC enforcement action will be effective in disciplining CEO compensation.

REFERENCES

Bebchuk, L. 2006. How much does the boss make? *Wall Street Journal* (January 18): A10.

BNA. 2007. Performance targets drew most comments in SEC review of statements, attorney says. *The Securities and Regulation Law Report* 39 (December 3): 46, 1826.

Byrd, J. W., M. F. Johnson, and S. L. Porter. 1998. Discretion in financial reporting: The voluntary disclosure of compensation peer groups in proxy statement performance graphs. *Contemporary Accounting Research* 15 (Spring): 25–52.

Chen, R., and R. Johnson. 2008. Securities and Exchange Commission comment letters: Enforcing accounting quality and disclosure. Working paper, The Ohio State University.

Core, J., W. Guay, and D. Larker. 2008. The power of the pen and executive compensation. *Journal of Financial Economics* 88 (1): 1–25.

Core, J., R. Holthausen, and D. Larker. 1999. Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics* 51 (3): 371–406.

Cox, C. 2006. *Speech by SEC Chairman: Remarks before the Council of Institutional Investors*. Available at: http://www.sec.gov/news/speech/spch033006cc.htm.

Cox, J., R. Thomas, and D. Kiku. 2003. SEC enforcement heuristics: An empirical inquiry. *Duke Law Journal* 53 (2): 737–779.

Cremers, M., and Y. Grinstein. 2010. The market for CEO talent: Implications for CEO compensation. Working paper, Yale University.

Dyck, A., N. Volchkova, and L. Zingales. 2008. The corporate governance role of the media: Evidence from Russia. *Journal of Finance* 63 (3): 1093–1135.

Dye, R. 2001. An evaluation of "essays on disclosure" and the disclosure literature in accounting. *Journal of Accounting and Economics* 32 (1–3): 181–235.

Ertimur, Y., F. Ferri, and V. Muslu. 2011. Shareholder activism and CEO pay. *Review of Financial Studies* 24 (2): 535–592.

Ettredge, M., D. Smith, and M. Stone. 2001. How effective is joint public and private monitoring? The case of the AICPA auditor change notification letter. *Auditing: A Journal of Practice & Theory* 20 (1): 29–44.

Ferri, F., and D. Maber. 2010. Say on pay votes and CEO compensation: Evidence from the U.K. Working paper, New York University.

Fischer, P., J. Gramlich, B. Miller, and H. White. 2009. Investor perceptions of board performance: Evidence from uncontested director elections. *Journal of Accounting and Economics* 48 (2–3): 172–189.

Grossman, S. 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24 (December): 461–483.

Harris, M. 1998. The association between competition and managers' business segment reporting decisions. *Journal of Accounting Research* 36 (1): 111–128.

Hartzell, J., and L. Starks. 2003. Institutional investors and executive compensation. *Journal of Finance* 58 (6): 2351–2374.

Heckman, J. J. 1978. Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46 (6): 931–959.

Holthausen, R. 2009. Accounting standards, financial reporting outcomes, and enforcement. *Journal of Accounting Research* 47 (2): 447–458.

Jackson, H., and M. Roe. 2009. Public and private enforcement of securities laws: Resource-based evidence. *Journal of Financial Economics* 93 (2): 207–238.

Jensen, M. 1979. Toward a theory of the press. In *Economics and Social Institutions*, edited by Dordrecht. Boston, MA: Martinus Nijhoff Publishing Company.

Jensen, M. 1993. The modern industrial revolution, exit and the failure of internal control systems. *Journal of Finance* 48 (3): 831–880.

Karuna, C. 2007. Industry product market competition and managerial incentives. *Journal of Accounting and Economics* 43 (2–3): 275–297.

Laksmana, I. 2008. Corporate board governance and voluntary disclosure of executive compensation practices. *Contemporary Accounting Research* 25 (4): 1147–1182.

Levitt, A. 1998. The importance of high quality accounting standards. *Accounting Horizons* 12 (1): 79–82.

Miller, G. 2006. The press as a watchdog for accounting fraud. *Journal of Accounting Research* 44 (5): 1001–1033.

Murphy, K., and J. Zabojnik. 2007. Managerial capital and the market for CEOs. Working paper, Queen's University.

Paredes, T. 2009. *Remarks before the Symposium on "The Past, Present, and Future of the SEC."* Available at: http://www.sec.gov/news/speech/2009/spch101609tap.htm.

Parratt, S. 2009. *Executive Compensation Disclosure: Observations on the 2009 Proxy Season and Expectations for 2010.* Available at: http://www.sec.gov/news/speech/2009/spch110909sp.htm.

Rajgopal, S., T. Shevlin, and V. Zamora. 2006. CEOs' outside employment opportunities and the lack of relative performance evaluation in compensation contracts. *Journal of Finance* 61 (4): 1813–1844.

Rock, S., S. Sedo, and M. Willenborg. 2001. Analyst following and count-data econometrics. *Journal of Accounting and Economics* 30 (3): 351–373.

Scannell, K. 2006. SEC steps up efforts to curb options dating in rule on executive pay. *The Wall Street Journal* (June 14): A3.

Scannell, K., and J. Lublin. 2006. Increased disclosure: Other perks will see light of day. *The Wall Street Journal* (June 27): C1.

Scannell, K., and J. Lublin. 2008. SEC unhappy with answers of executive pay. *Wall Street Journal* (January 29): B1.

Schwartz, K., and B. Soo. 1996. Evidence of regulatory noncompliance with SEC disclosure rules on auditor changes. *The Accounting Review* 71 (4): 555–572.

Stigler, G. 1970. The optimum enforcement of laws. *The Journal of Political Economy* 78 (3): 526–536.

Verrecchia, R. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5 (1): 179–194.

Verrecchia, R. 2001. Essays on disclosure. *Journal of Accounting and Economics* 32 (1–3): 97–180.

White, J. 2007. *Speech by SEC Staff: Where's the Analysis?* Available at: http://www.sec.gov/news/speech/2007/spch100907jww.htm.

APPENDIX A

Variable Definitions

NDEFECT = the total number of noncompliance comment items in the SEC's critique;

PDEFECT = the number of noncompliance comment items in the SEC's critique that are related to the firm's disclosures on the content of management compensation;

GDEFECT = the number of noncompliance comment items in the SEC's critique that are related to the firm's disclosures on corporate governance mechanisms;

RDEFECT = the number of noncompliance comment items in the SEC's critique that are related to the readability of the firm's compensation disclosures;

EXPECTCOMP = the expected compensation for the CEO, estimated using the predicted value from a regression of log-transformed total CEO compensation (in thousands) on the economic determinants following Core et al. (2008) (see Appendix B);

EXCESSCOMP = the excess compensation for the CEO, estimated using the residuals from a regression of log-transformed total CEO compensation (in thousands) on the economic determinants following Core et al. (2008) (see Appendix B);

DIFF = the extent of product differentiation in a firm's industry (four-digit SIC code), measured as sales divided by operating costs (costs of goods sold, SGA expenses, depreciation, depletion, and amortization) for the industry, following Karuna (2007);

Log(*INDSZ*) = the logarithm of the market size in a firm's industry (four-digit SIC code), measured following Karuna (2007) as total industry sales (in millions);

Log(*ENTRY*) = the logarithm of entry cost in a firm's industry (four-digit SIC code), measured as weighted average of property, plant and equipment (in millions) for a firm's industry (the weight is each firm's market share in the industry);

NPRESS = indicator equal to 1 for firms with at least one press article criticizing the firms' executive compensation in the year;

LPRXY = the logarithm of the total number of words in a firm's proxy statement for the year;

RET = buy-and-hold stock returns during the year;

ROA = return on assets (income before extraordinary items divided by average total assets) for the year;

LSZ = the logarithm of the market capitalization (in millions) at the year-end;

CHAIR = indicator equal to 1 if CEO is also the chairman of the board of directors;

INDIR = the percentage of the directors who are "independent," defined as a director with "no material relationship" with the listed company, either directly or as a partner, shareholder, or officer of an organization that has a relationship with the company;

DIRCEO = the percentage of the directors who are outside directors appointed by CEO;

DIRGRAY = the percentage of the directors who are gray outside directors (a director is "gray" if she or her employer receive payments from the company in excess of her board pay);

DIRBUSY = the percentage of the directors who are busy outside directors (a director is "busy" if she serves on three or more other boards (six or more if she is retired);

LBOARD = the logarithms of the number of directors on the board; and

INST5 = the percentage of total shares outstanding held by the top five institutional investors.

APPENDIX B

Methodology for Estimating the Expected and Excess Components of CEO Compensation

This appendix describes the method we use to estimate the amount of CEO compensation in excess of the pay justified by the economic determinants of CEO compensation. We follow Core et al. (2008) and estimate an OLS regression of CEO compensation on the economic determinants of CEO compensation as follows:

$$\begin{aligned} Log(Compensation_i) = {} & \alpha_0 + \alpha_1 Log(Tenure_i) + \alpha_2 S\&P500_i + \alpha_3 Log(Sales_i) + \alpha_4 BM_i \\ & + \alpha_5 RET_i + \alpha_6 LagRET_i + \alpha_7 ROA_i + \alpha_8 LagROA_i + u_i, \end{aligned} \tag{1}$$

where:

Log(*Compensation*) = the logarithm of total compensation for CEO (i.e., salary, bonus, long-term incentive plan payouts, the value of restricted stock grants, the value of options granted during the year, and any other annual pay for the CEO) for fiscal year 2006, and the remaining variables consist of the following economic determinants:

Log(*Tenure*) = the logarithm of the CEO's tenure (in years) at the end of 2006;
S&P500 = indicator variable equal to 1 for firms in the S&P500 index at the end of 2006, and 0 otherwise;
Log(*Sales*) = logarithm of the firm's sales in 2005;
BM = book-to-market ratio measured at the end of 2005;
RET = firm's buy-and-hold return for 2006;
LagRET = firm's buy-and-hold return for 2005;
ROA = return on assets (income before extraordinary items divided by average total assets) for 2006; and
LagROA = return on assets for 2005.

We also include industry fixed effects (two-digit SIC code) in the regression to control for industry effects.[24]

The results of estimating Equation (1) with the sample of firms critiqued by the SEC are reported in Table B1. We find that this determinant model exhibits great explanatory power for the CEO compensation: it explains 60 percent of the variation in the CEO compensation. Turning to the coefficients of the independent variables, we find that the compensation is negatively associated with firm size, growth opportunity, and stock performance. These results are consistent with prior research (e.g., Core et al. 2008).

We then separate 2006 CEO compensation into two components: the prediction from the estimated CEO compensation regression (expected pay) and a residual (excess pay). Specifically, we compute excess compensation by estimating expected compensation for fiscal year 2006 from Equation (1) and subtracting it from actual compensation in 2006.

---

[24] Core et al. (1999) suggest that industry indicators control for the differences in the demand for CEO talent. In untabulated analysis, we find our results are robust to using two alternative proxies for CEO outside opportunities from Rajgopal et al. (2006)—a three-year industry-adjusted return on assets and the number of articles citing the CEO in major U.S. and global business newspapers and wire services.

TABLE B1

Regression Results for Decomposing Total Compensation into Expected Compensation and Excess Compensation

Independent Variables	**Dependent Variable: *Log(Compensation)***
Log(Tenure)	0.031 (0.051)
S&P500	0.317** (0.132)
Log(Sales)	0.357*** (0.050)
BM	−0.052 (0.236)
RET	0.650*** (0.187)
LagRET	0.418** (0.176)
ROA	−0.690 (0.978)
LagROA	−0.158 (0.818)
Industry fixed effects	Included
R^2	0.60

, * Indicate two-tailed statistical significance at 5 percent and 1 percent levels, respectively.

Variable Definitions:

Log(Total Compensation) = the logarithm of salary, bonus, long-term incentive plan payouts, the value of restricted stock grants, the value of options granted during the year, and any other annual pay for the CEO in 2006;
Log(CEO Tenure) = the logarithm of the CEO's tenure (in years) at the end of 2006;
S&P500 = an indicator variable equal to 1 for firms in the S&P500 index at the end of 2006, and 0 otherwise;
Log(Sales) = the logarithm of the firm's sales in 2005;
BM = the book-to-market ratio measured at the end of 2005;
RET = the firm's buy-and-hold return for 2006;
LagRET = the firm's buy-and-hold return for 2005;
ROA = the firm's return on assets (income before extraordinary items divided by average total assets) for 2006; and
LagROA = the firm's return on assets for 2005.

This table reports the results of OLS regression of the logarithms of CEOs' total compensation on the economic determinants of compensation. The regression model follows Core et al. (2008). The sample consists of 336 firms whose compensation disclosures in the proxy statements of fiscal year 2006 were reviewed by the SEC.

Industry fixed effects are measured using two-digit SIC codes. The standard errors are reported in parentheses.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1445–1475

American Accounting Association
DOI: 10.2308/accr-10032

Data Truncation Bias, Loss Firms, and Accounting Anomalies

Siew Hong Teoh
University of California, Irvine

Yinglei Zhang
Chinese University of Hong Kong

ABSTRACT: *Ex post* trimming of extreme returns observations that are not data errors causes spurious inferences in tests of market efficiency and behavioral explanations for anomalies. Trimming causes a downward truncation bias in estimated mean returns that is stronger in *ex ante* subsamples with more loss firms and in which return distributions are more right-skewed. There is an asymmetric U-shaped relation between return right-skewness and loss frequency across deciles of negative return predictors (*Accruals*, *ΔNOA*, and *NOA*), and a downward sloping relationship for positive return predictors (*CFO* and *FCF*). Consequently, a least-trimmed square (LTS) 1 percent deletion of returns induces a spurious inverted-U-shaped relation between returns and negative predictors, and an exaggerated positive relation for positive predictors. Thus, the resulting trimmed relations do not reject behavioral explanations for these anomalies. Trimming also induces a spurious loss anomaly. These findings highlight that in return prediction studies, observations should not be deleted based upon the values of the dependent variable, only based upon clearly identified data errors.

Keywords: *truncation bias; data trimming; anomalies; market efficiency; behavioral finance; accruals; net operating assets; cash flows from operations; free cash flows.*

JEL Classifications: *C24; C58; G14; M41.*

I. INTRODUCTION

An important topic for investment professionals, regulators, and accounting scholars in the empirical capital markets area is whether accounting-based variables predict abnormal returns. After controlling for standard risk factors, accrual-related variables are found to

---

We thank workshop and conference participants at the University of California, Irvine, Chinese University of Hong Kong, Nanyang Business School at Nanyang Technological University, Financial Accounting Research Section (FARS) 2006 Conference, and 2006 American Accounting Association Annual Meeting. We also thank David Hirshleifer, Joanna Ho, Bin Ke, Mort Pincus, Anne Beatty, Doug Schroeder, Charles Wasley (FARS conference discussant), Scott Richardson (the AAA discussant), and Huai Zhang for valuable comments. The authors are grateful for able research assistance from Xuan Huang, Peng-Chia Chiu, Karen Lin, and Chunquan Zhou.

Editor's note: Accepted by Paul Zarowin.

Submitted: January 2010
Accepted: January 2011
Published Online: April 2011

predict abnormal future returns negatively, whereas cash-flow-related variables are found to predict abnormal future returns positively; see Sloan (1996) for the accruals anomaly (*Accruals*), Fairfield et al. (2003) and Richardson et al. (2005) for the change in net operating assets anomaly (ΔNOA), Hirshleifer et al. (2004) for the net operating assets anomaly (*NOA*), Desai et al. (2004) for the operating cash flows anomaly (*CFO*), and Hackel et al. (2000) for the free cash flow anomaly (*FCF*). Generally, these studies conclude that the accounting-based anomalies reflect investor behavioral biases.[1] In this view, investors fail to incorporate fully the information contained in financial reports due to functional fixation on earnings or limited attention.

Kraft et al. (2006; hereafter, KLW) challenge the behavioral explanation for accounting anomalies. They report that these anomalies are not robust to *ex post* removal of extreme observations. They conduct a Least-Trimmed Squared regression (hereafter, LTS) at 1 percent by deleting those observations whose regression residuals are among the largest 1 percent. After trimming, they find an inverted-U relation between *Accruals* or *NOA* and subsequent one-year buy-and-hold size-adjusted returns. They argue that the behavioral theory predicts a monotonic relation between future returns and *Accruals* or *NOA* that should not be driven by just a small fraction of the observations, leading them to conclude that the non-robustness of monotonicity to trimming raises a severe challenge to behavioral explanations of these anomalies.

The presumption that an empirical finding should be robust to the trimming of a small fraction of extreme observations sounds plausible. One possible justification is that extreme observations are more likely to be data errors. However, right-skewed distributions tend to generate legitimate extreme positive observations. In robustness tests of psychological biases or rational market pricing, *ex post* truncation of such observations biases the estimated trimmed mean return downward. We refer to this phenomenon as *truncation bias*, which is a special case of the familiar statistical concept of selection bias. Researchers therefore must be cautious in interpreting evidence obtained after trimming of extreme observations when the return distribution is skewed.

Since the publication of KLW, the literature has largely taken at face value their inference that behavioral explanations for accounting anomalies fail. Furthermore, it has become common to use trimming of extreme observations for primary analyses, and not just to verify robustness; readers are often not even informed about the results of the untrimmed population. A search of ISI Web of Science, Google, and SSRN in May 2010 yielded 50 articles referencing KLW.[2] With few exceptions, these studies do *not* alert the reader to the truncation bias introduced by the LTS procedure. Many cite KLW uncritically as the reason to use the same selection procedure (either trimming observations or winsorizing the dependent variable returns) and/or as refuting behavioral explanations for anomalies.

In this study, we demonstrate that *ex post* LTS trimming of extreme returns that are not data errors is an incorrect procedure for testing robustness of either market efficiency or anomaly studies, and that KLW's inference that behavioral theories fail to explain the anomalies is mistaken. We consider the effects of the induced truncation bias from LTS trimming on a set of accrual-related anomalies (*Accruals* and *NOA*, which were considered by KLW, and ΔNOA), and cash-flow-related anomalies (*CFO* and *FCF*).

---

[1] A branch of the literature on the accrual anomaly focuses on earning management to take advantage of investor naiveté and influence the behavior of auditors, analysts, and investors; e.g., Teoh et al. (1998a, 1998b), Ali et al. (2000), Bradshaw et al. (2000), Xie (2001), Teoh and Wong (2002), Collins et al. (2003), Fairfield et al. (2003), Dechow and Ge (2006), Dechow et al. (2008), Ali et al. (2008), Levi (2008), and Louis et al. (2008).

[2] Twenty-nine of these are either published or forthcoming, and 16 are in the top 5 accounting journals (3 in *The Accounting Review*, 3 in *Journal of Accounting and Economics*, 4 in *Journal of Accounting Research*, 3 in *Contemporary Accounting Research*, and 3 in *Review of Accounting Studies*). The remaining 21 articles are working papers.

Stock return distributions tend to be right-skewed, such that LTS deletion tends to remove large, positive rather than negative abnormal returns. The larger the right-skewness of returns for a decile portfolio, the larger is the downward truncation bias for that decile from LTS trimming. When right-skewness of returns varies across predictor deciles, the induced truncation bias from LTS trimming across predictor deciles will tend to follow the pattern across deciles in skewness. We provide evidence that the shape of the truncation bias resembles the shape of the distribution of returns' right-skewness and of the prevalence of losses across predictor deciles. Extreme accounting predictor deciles have higher right returns skewness and higher prevalence of losses than the middle deciles, so LTS deletion induces a greater truncation bias at the extreme portfolios, by as much as three to four times, than in the middle deciles. In other words, the truncation bias has an inverted U-shape. It is also asymmetric. The lowest decile experiences the highest right skew and prevalence of losses and, hence, the largest downward truncation bias. The truncation bias is at least twice as large in the lowest decile than in the highest decile.

We also find that the truncation bias is large in magnitude, such that the inverted U-shape of the bias across predictor deciles swamps the original monotonic negative relation between returns and accrual-related predictors to create the inverted U-shape pattern identified by KLW. In contrast, for cash-flow-related anomalies, the similarly asymmetric inverted U-shaped truncation bias exaggerates the positive slope of the relation between future returns and *CFO* or *FCF*, such that these anomalies appear strengthened after trimming, and monotonicity of these relations is preserved. These contrasting results would cause a researcher who neglects truncation bias to infer incorrectly that behavioral explanations fail for accrual-related anomalies but hold for cash-flow-related anomalies.

KLW's message has resonated with researchers because it seems plausible that anomalies lack robustness if deleting a mere 1 percent of the sample can overturn monotonicity. Indeed, KLW (2006, 301) assert that "for 99 percent of the ... sample, there is an inverted U-shaped relation between accruals and returns. An inverted U-shaped relation is inconsistent with the earnings fixation hypothesis offered by Sloan (1996), which predicts a monotonically decreasing relation between BHARs and accruals." However, *ex post* trimming even a small fraction of a sample can lead to large spurious effects because, in general, return-predicting variables explain only a small fraction of the variation in returns (low R^2).[3] There is therefore a strong *ex ante* reason to expect that deleting even a small fraction of observations based upon their having extreme returns can have effects that are quite large relative to the relation being estimated.

Nevertheless, owing to the intuitive appeal of the idea that a small fraction of observations should not be influential, many accounting scholars will place greater credence upon behavioral explanations for anomalies if it can be shown that the monotonic relation for anomalies is not driven by a small fraction of the observations. Thus, we devise two alternative ways of implementing trimming to see whether, in contrast with KLW, monotonicity of anomalies survives trimming. In the first approach, we perform a uniform 1 percent LTS trimming within each predictor decile. While this does not eliminate the truncation bias, uniform 1 percent trimming reduces the severity of downward truncation bias at the extreme deciles in which right-skewness is the highest. We find that within-decile 1 percent LTS trimming preserves the monotonic shape of the relation between returns and predictor variables in the remaining 99 percent of the sample.

Our second approach identifies *ex ante* firm characteristics that are associated with *ex post* high right-skewness of returns across predictor deciles. By eliminating observations based on *ex ante* characteristics associated with high skewness, the variation in *ex post* returns' right-skewness across predictors can be reduced. This makes it less likely that the truncation bias from LTS trimming within the selected sample will alter the shape of the original relation between predictors and returns.[4]

---

[3] This has to be the case if markets have even a minimal degree of efficiency. If a variable could predict future returns with a high R^2, then it would provide an enormous low-risk money machine. Arbitrageurs have strong incentives to exploit such opportunities until predictive power is reduced to the point where arbitrage opportunities become riskier (Shevlin et al. 2006).

A challenge for this approach is that firm characteristics that are associated with right returns' skewness can also be associated with drivers of the anomaly. For example, small firms with highly volatile returns have highly positive skewed returns, but prior literature has found that anomalies are stronger in these firms.[5] Therefore, eliminating such firms from the test sample could "throw the baby out with the bathwater" by removing key sources of mispricing. The fraction of loss firms is associated with high right returns skewness,[6] but in contrast with size or high volatility, past literature indicates that anomalies tend to be *weaker* among loss firms. Consequently, eliminating such firms can reduce the variation in truncation bias among anomaly deciles without significantly weakening the strength of the anomaly.[7]

We therefore remove loss firms from the sample, and explore the effects of trimming on the resulting 79 percent of the sample, identified *ex ante*, that are profit-only firms. We find that the truncation bias from trimming in the profit-only sample is smaller than in the full sample. As a result, the monotonic relation in returns across predictor deciles is preserved in the trimmed profit-only sample. While *ex post* trimming is still inappropriate, the robustness of the pattern of anomalies in large subsamples (99 percent with uniform trimming across deciles, and 79 percent of profit firms) will give confidence to researchers that the monotonic relation of anomalies is indeed robust.

Furthermore, we find that the use of more appropriate asset-pricing methods and controls significantly reduce the effects of trimming on monotonicity. Although the use of appropriate methods is important for the test for anomalies (Zach 2003), KLW control only for size, and use pooled OLS regressions and annual horizon abnormal returns. Following convention in the asset-pricing and accounting literatures (Fama and French 1992; Hirshleifer et al. 2004), we use monthly returns and standard asset-pricing controls beyond just size, including book-to-market, 12-month momentum, and past one-month and three-year returns (the short-term and long-term reversal effects), as well as the Fama and MacBeth (1973; hereafter, Fama-MacBeth) procedure. Monthly returns are less right-skewed than annual returns, and residuals are statistically better behaved when risk is more appropriately controlled. The monthly cross-sectional regressions in the Fama-Macbeth procedure also minimize survivorship bias, and the time-series t-statistics correct for cross-sectional dependence of residuals. We demonstrate that, with better behaved return residuals, the induced truncation bias from trimming is less severe.

Two prior studies have also raised concerns about *ex post* trimming of observations. Core (2006) warns that non-random deletion of observations can potentially lead to spurious inferences in contexts other than KLW's. Nevertheless, he suggests that "KLW's robustness tests may be appropriate in their specific setting" (Core 2006, 350). We provide evidence and explain why non-random trimming of extreme observations induces bias in robustness tests of a comprehensive set of five accounting anomalies. Our contribution goes beyond Core's (2006) discussion in several ways. We identify the pattern of the truncation bias across predictor variables and show that losses can be used to identify how truncation bias varies across predictor deciles. We show that the truncation bias strengthens cash-flow-related anomalies. We offer *ex ante* truncation methods, and return measurement and regression techniques that reduce right returns skewness, and therefore reduce truncation bias. Finally, we provide evidence that KLW's claim of non-robustness of the monotonicity of anomalies to trimming is itself not robust.

---

[4] Trading strategies for anomalies are based on top and bottom predictor portfolios so the magnitude of hedge profits can be robust to *ex post* trimming of returns if the induced truncation bias is approximately symmetric for the top and bottom portfolios even when the truncation bias has a U-shape with the predictor deciles.

[5] Such firms are more likely to have poorer information environments and so are more likely to be mispriced, and the mispricing is more costly to arbitrage.

[6] Intuitively, since losing money is not viable in the long-term, loss firms are likely to engage in strategic business shifts (possibly associated with financial distress, high leverage, and restructuring). In consequence, we expect the future earnings and returns of loss firms to be much more variable than those of profit firms. When downside stock returns are bounded below by −100 percent but not bounded above, the higher dispersion in profitability induces greater positive skewness. See footnote 15 for references that anomalies are weaker for loss firms.

[7] Furthermore, unlike size and volatility, the distribution of loss firms across predictor deciles is similar to the actual observed asymmetric and inverted U-shaped distribution of the truncation bias. Loss firms have greater right-skewness in returns and are most concentrated in the lowest deciles.

Kothari et al. (2005) warn that non-random deletion can lead to a spurious finding of return predictability. Their point does not directly oppose KLW's inference *against* the robustness of predictability. It is, however, consistent with our message that non-random deletion is inappropriate as a robustness test because it induces a truncation bias. Kothari et al. (2005, 152) raise the conceptual concern about spurious predictability using only simulated data, stating that "our simulations do not address whether the magnitude of the biases we report can be seen in actual data." We test whether trimming causes a spurious rejection of market efficiency using actual data. We find that a loss indicator variable does not predict abnormal returns in a full sample, but that there is spurious return predictability in a sample trimmed using the 1 percent LTS procedure. This evidence together with the evidence on trimming in anomalies studies show that non-random deletion is inappropriate for testing robustness of either market efficiency or anomalies.

Section II describes the sample selection and defines the empirical variables. Section III discusses how truncation bias from *ex post* trimming of extreme returns induces a spurious inverted-U relation. It also explores the effects of abnormal return measure, regression method, and inadequate asset-pricing controls on the truncation bias. Section IV presents results from LTS trimming in each predictor decile and in the profit subsample. It also investigates the relevance of the distribution of losses across predictor deciles for truncation bias, and whether trimming induces spurious rejection of market efficiency. Section V concludes.

II. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

Sample Selection

The sample covers NYSE, AMEX, and NASDAQ firms with available data from July 1964 to December 2002, from the intersection of CRSP and Compustat datasets. All observations are required to have sufficient financial data to compute net operating assets, accruals, earnings, firm size, the book-to-market ratio, and 11 months of past returns. The last three variables are needed to compute the Daniel et al. (1997) three-factor characteristics adjusted abnormal returns.

KLW discuss several possible selection biases in anomalies research. First, most studies limit analyses to NYSE/AMEX firms and may depend on either ZLIST (from Compustat) or HEXCD (from CRSP) to identify the exchange listing. However, only the EXCHCD variable (from CRSP) provides the correct historical identification. They document that this subtle choice can affect the documented *Accruals* hedge returns. Second, the exclusion of firm-year observations in which year $t+1$ *Earnings* or *Accruals* are missing, which is imposed by some previous studies in order to perform the Mishkin (1983) test, can introduce a survivorship bias. A further selection bias involves excluding firm-year observations with missing returns during the BHAR accumulation period.

We use a procedure that avoids all three potential sample selection biases. We include firms on NASDAQ as well NYSE/AMEX, such that there is no misidentification problem across the exchange listing. The results here do not rely on the Mishkin (1983) test, so no constraint for the availability of future financial data is imposed in our test design. Finally, the portfolios examined in this study are rebalanced monthly to avoid survivorship bias and measurement problems for long-term abnormal returns highlighted by Fama (1998).[8]

Financial Variable Definition and Measurement

Following Hirshleifer et al. (2004), we calculate net operating assets as the difference between operating assets and operating liabilities, where:

$$\begin{aligned}&\mathit{Operating\ Assets}_t = \mathit{Total\ Assets}_t - \mathit{Cash\ and\ Short\text{-}Term\ Investment}_t;\\&\mathit{Operating\ Liabilities}_t = \mathit{Total\ Assets}_t - \mathit{Short\text{-}Term\ Debt}_t - \mathit{Long\text{-}Term\ Debt}_t\\&\qquad - \mathit{Minority\ Interest}_t - \mathit{Preferred\ Stock}_t - \mathit{Common\ Equity}_t. \qquad (1)\end{aligned}$$

We calculate operating accruals using the indirect balance sheet approach (Sloan 1996) for all fiscal years in the sample, since the data coverage starts before 1987, when SFAS 95 was effective and cash flow from operations numbers are directly reported and available from Compustat (Hribar and Collins 2002).

$$\begin{aligned}\mathit{Accruals}_{bs} = {}& (\Delta \mathit{Current\ Assets} - \Delta \mathit{Cash\ and\ Short\text{-}Term\ Investment})\\&- (\Delta \mathit{Current\ Liabilities} - \Delta \mathit{Short\text{-}Term\ Debts} - \Delta \mathit{Tax\ Payable})\\&- \mathit{Depreciation}. \qquad (2)\end{aligned}$$

$$\mathit{Cash\ Flow\ from\ Operations} = \mathit{Operating\ Earnings} - \mathit{Accruals}_{bs}. \qquad (3)$$

$$\mathit{Free\ Cash\ Flows} = \mathit{CFO} - \mathit{Capital\ Expenditures}. \qquad (4)$$

The key predictor variables are operating accruals, change in net operating assets, net operating assets, cash flow from operations, and free cash flows, all scaled by the fiscal year beginning total assets.[9] To avoid the extremely small denominator problem, we delete observations with lagged total assets less than $1,000,000. If short-term debt, tax payable, long-term debt, minority interest, or preferred stock have missing values, then we treat these values as zeros to avoid unnecessary loss of observations. Table 1 summarizes the calculation to obtain the five predictor variables, *Accruals*, *ΔNOA*, *NOA*, *CFO*, and *FCF*, and their summary statistics, which are in a similar range as those in the prior literature. In addition, an indicator variable *DLOSS* takes the value 1 if the firm incurred a loss in the fiscal year, and 0 otherwise. Table 1 shows that an average of 21.6 percent of firm-year observations incurred losses during the sample period.

We use the 11-month intermediate return horizon starting 12 months through 2 months prior to the portfolio formation month as the measure of price momentum. Following the convention in the finance literature, for this variable we skip the month just prior to the testing month (for which we control separately).

Calculation of Abnormal Returns

When measuring abnormal portfolio returns for computing profits from the trading strategy, we adopt the characteristics-adjustment approach proposed by Daniel et al. (1997) to control

---

[8] For several reasons (including non-normality and bad model problems), Fama (1998, 291) maintains that monthly returns are superior to longer horizon returns for asset-pricing tests. He asserts that "formal inferences about long-term returns should be based on averages or sums of short-term abnormal returns (AARs or CARs) rather than ... buy-and-hold abnormal returns (BHARs)" because the "bad-model errors in expected returns grow faster with the return horizon than the volatility of returns" (Fama 1998, 285).

[9] The main results are not affected by the choice of ending total assets or average total assets as deflators.

TABLE 1
Descriptive Statistics

	Accruals	*NOA*	*ΔNOA*	*CFO*	*FCF*	*DLOSS*
Mean	−0.0160	0.7672	0.1155	0.0665	−0.0483	0.2161
Median	−0.0307	0.7260	0.0531	0.1169	0.0355	0.0000
Std. Dev.	0.2652	1.0009	0.5875	0.4836	0.6698	0.4116
P1	−0.4223	−0.0612	−0.4367	−1.1977	−1.6634	0.0000
P99	0.5821	2.4970	1.4426	0.5889	0.48678	1.0000

Variable Definitions:
Net Operating Assets = Operating Assets (OA) − Operating Liabilities (OL), where:
  OA = total assets − cash and short-term investment;
  OL = total assets − STD − LTD − MI − PS − CE, where:
    STD = debts included in current liabilities;
    LTD = long-term debts;
    MI = minority interests;
    PS = preferred stocks;
    CE = common equity;
  NOA = net operating assets/lagged total assets;
  ΔNOA = change in net operating assets/lagged total assets;
Operating Accruals = (ΔCA − $\Delta Cash$) − (ΔCL − ΔSTD − ΔTP) − Dep, where Δ is the annual change in;
  CA = current assets;
  CL = current liabilities;
  TP = income tax payable; and
  Dep = depreciation and amortization;
$Accruals$ = operating accruals/lagged total assets;
CFO = (earnings before extraordinary item − operating accruals)/lagged total assets;
FCF = CFO − capital expenditure/lagged total assets; and
$LOSS$ = 1 if earnings before extraordinary item < 0; otherwise 0.

CUT LIBRAR UNIVERS

simultaneously for the following asset-pricing characteristics: size, book-to-market (B/M), and momentum. Daniel and Titman (1997) find evidence that characteristics provide better *ex ante* cross-sectional forecasts of future returns than do factor loadings.

Specifically, each month, we form size-B/M-momentum benchmark portfolios ($5 \times 5 \times 5 = 125$) based on NYSE cutoff points. To calculate monthly characteristics-adjusted returns (hereafter, abnormal returns), we then subtract the equal-weighted returns of the size-B/M-momentum benchmark portfolio to which the firm belongs from the raw buy-and-hold returns for each firm each month. To ensure the availability of the financial information, we match returns with the financial data starting from the 5th month to the 16th month after fiscal year-end.

III. UNDERSTANDING TRUNCATION BIAS FROM TRIMMING EXTREME RETURNS FIRMS

Ex post trimming of extreme observations from a right-skewed return distribution induces a downward truncation bias in the mean returns of the trimmed sample. This is because the return distribution is left bounded at −100 percent and truncation based on the least-trimmed squares procedure predominantly removes the right tail and not as much of the left tail extreme observations.

As an illustration to aid understanding, consider the truncation of extreme positive returns for the *Accruals* and *CFO* anomalies in Figures 1 and 2; the effects for *ΔNOA* and *NOA* are similar to

Accruals, and the effects for *FCF* are similar to *CFO*. Suppose that the return skewness does not vary cross-sectionally with predictors and the same return cutoff value is applied to truncate extreme positive return observations in every predictor decile. Then, the portfolio with the highest density of large positive returns (e.g., lowest *Accruals* or *CFO* deciles) will have the largest fraction of extreme positive returns eliminated and will therefore experience the largest downward bias in returns. On the other hand, the portfolio with the lowest density of positive returns (e.g., highest *Accruals* or *CFO* deciles) will have the smallest fraction of positive returns eliminated, and will experience the lowest downward bias in returns.

When the downward truncation bias in returns is constant across every decile, monotonicity of the relation between predictors and future returns is preserved, such that trimming neither affects the magnitude of the trading strategy profits nor biases inferences. When the downward truncation bias varies linearly across predictor deciles, trimming may weaken, eliminate, or reverse the anomaly, depending on the strength of the truncation bias relative to the underlying anomaly. Figure 1 illustrates the case where the downward truncation bias is decreasing with the predictor variable. The resulting effect from trimming is a weakened accrual anomaly and a strengthened *CFO* anomaly, but monotonicity is preserved for either anomaly.

Uneven truncation of return distributions by itself does not induce the inverse U-relation for the *Accruals* anomaly in KLW. The truncation bias needs to vary *nonlinearly*. Figure 2 illustrates a nonlinear truncation bias similar to actual evidence presented later where the bias is largest for the lowest *Accruals* or *CFO* deciles, decreases steeply and monotonically up to decile 5, and then increases slightly beyond decile 5. The downward truncation bias for the lowest *Accruals* decile is large enough to swamp the positive returns, thereby causing the net returns after trimming to be negative. The smaller truncation bias for the higher deciles exacerbates the negative returns. The resulting effect is an inverted U-shape pattern, as observed in KLW.

Figure 2, Panel B illustrates the effect of the nonlinear truncation bias from trimming for the *CFO* anomaly. The downward truncation bias adds to the untrimmed negative returns in the lowest deciles, whereas it reduces the positive returns in the highest deciles. The resulting effect from trimming is also an inverted U-shape, but monotonicity is preserved because the assumed downward bias is declining with *CFO* decile. The trading profits are actually larger after trimming. Thus, had KLW also evaluated LTS trimming on *CFO*-related anomalies, they would have concluded in favor of behavioral explanations for these type of anomalies. We turn next to examining the actual magnitudes of truncation bias by LTS trimming.

Monthly Returns Skewness and Truncation Bias by Predictor Deciles

Table 2 presents the right-skewness of monthly returns across predictor deciles for the five anomaly variables for the full sample in Panel A. We defer discussion of the evidence for the profit sample in Panel B to Section IV. Consistently across all predictor variables, the skewness of returns is largest for the lowest deciles, as compared to the highest deciles. For example, the lowest decile (L) has a higher return skewness than the highest decile (H) by 0.75 for *Accruals*, 1.10 for *ΔNOA*, 0.84 for *NOA*, 0.89 for *CFO*, and 0.58 for *FCF*. Viewed another way, the lowest decile return skewness relative to the highest decile skewness is higher by 50 percent for *Accruals*, almost double for *ΔNOA*, 66 percent for *NOA*, almost 80 percent for *CFO*, and almost 50 percent for *FCF*. As a result, the magnitudes of the truncation bias also vary considerably across deciles.

Table 3 reports the number of observations and magnitude of monthly returns that are deleted by LTS 1 percent trimming in each decile, separately grouped by positive abnormal returns observations versus negative abnormal returns observations. Many more extreme positive abnormal return observations are deleted than extreme negative observations, by an order of magnitude larger, across all five accounting predictors. It is also remarkable that the number of observations and

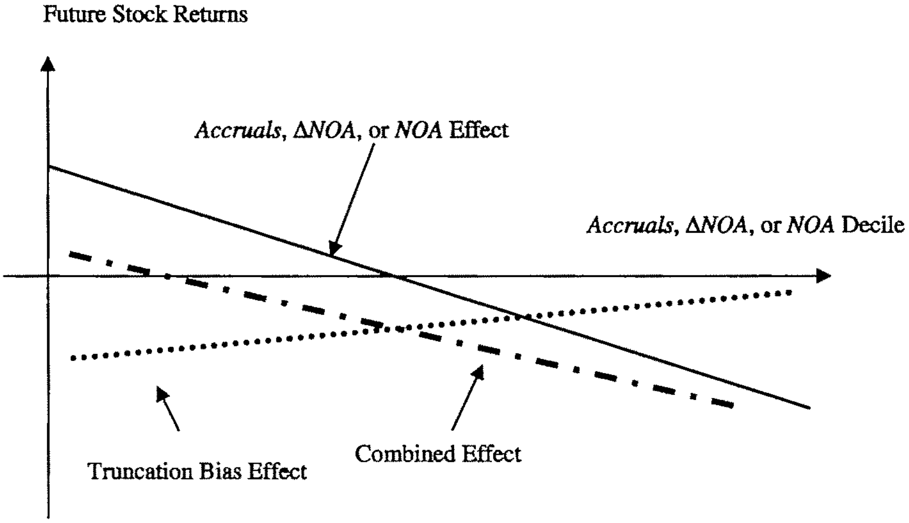
FIGURE 1

Linear Truncation Bias on Accounting Anomalies

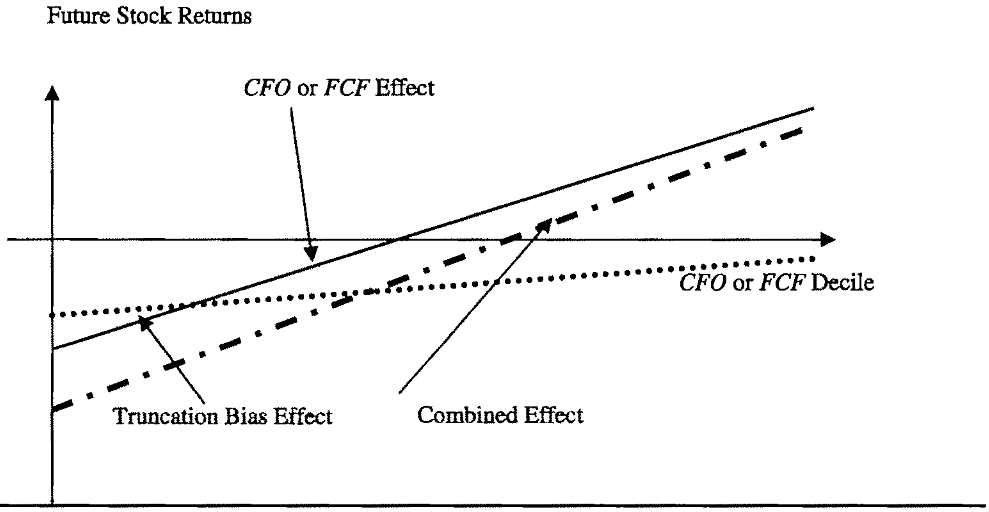
Panel A: The Effect of Truncation Bias on the *Accruals*, Δ*NOA*, or *NOA* Anomaly if Truncation Bias is Linear

Panel B: The Effect of Truncation Bias on the *CFO* or *FCF* Anomaly if Truncation Bias is Linear

Panel A


Future Stock Returns
Accruals, Δ*NOA*, or *NOA* Effect
Accruals, Δ*NOA*, or *NOA* Decile
Truncation Bias Effect
Combined Effect

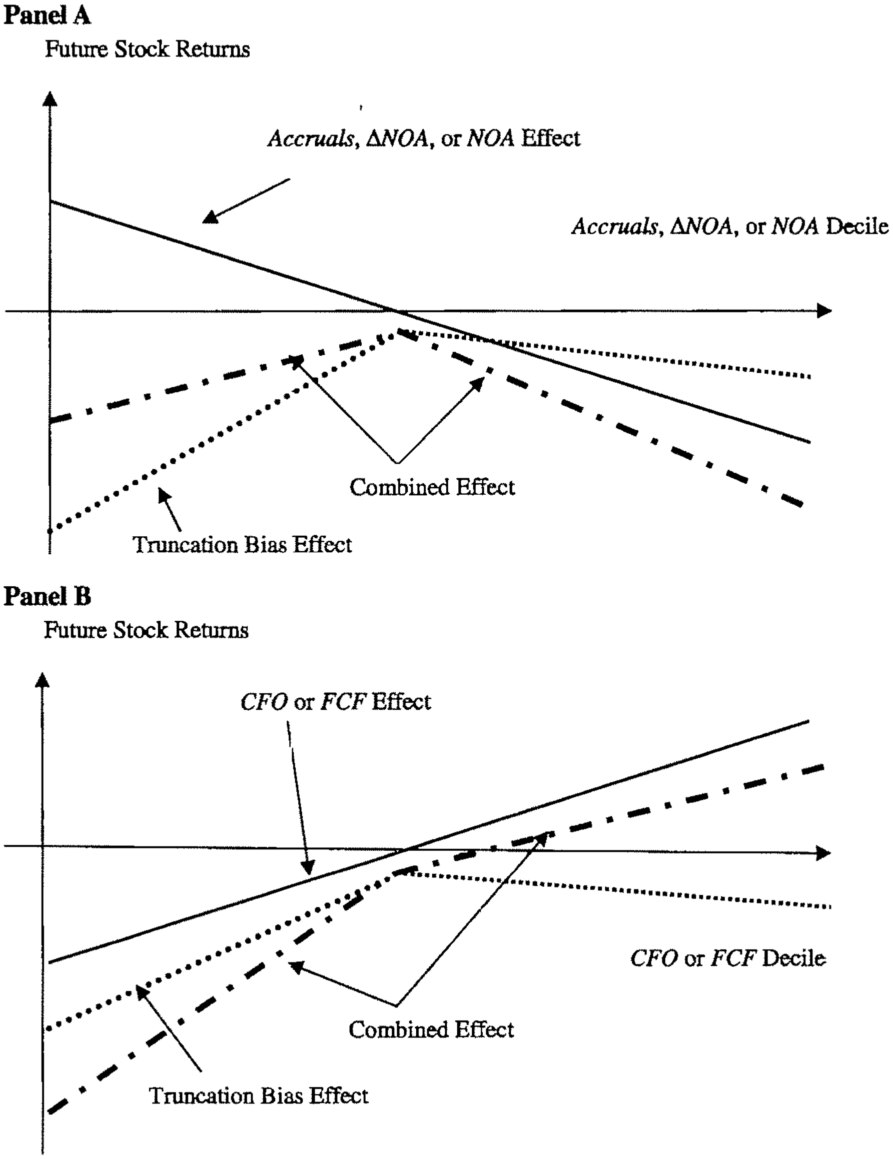
Panel B



magnitude of returns deleted exhibit an approximate inverted U-shape across deciles for both the positive and negative extreme observations. The extreme deciles suffer more deletions from LTS trimming than the middle deciles, and the lowest decile consistently suffers the most when compared to other deciles.

FIGURE 2
Nonlinear and Inverted U-Shaped Truncation Bias on Accounting Anomalies

Panel A: The Effect of Truncation Bias on the *Accruals*, *ΔNOA*, or *NOA* Anomaly if the Truncation Bias is Nonlinear and Inverted U-Shaped
Panel B: The Effect of Truncation Bias on the *CFO* or *FCF* Anomaly if the Truncation Bias is Nonlinear and Inverted U-Shaped



Because the number of extreme positive return observations dominates the number of extreme negative return observations, the net effect of trimming is a systematic inverted U-shape across the deciles. For example, Panel A of Table 3 shows that much larger positive returns are trimmed from both the lowest and the highest *Accruals* deciles than from the middle groups. For the lowest decile,

TABLE 2
Mean Return Skewness across Accounting Predictor Deciles

Panel A: All-Firm Sample

Decile	*Accruals*	*ΔNOA*	*NOA*	*CFO*	*FCF*
L	2.176	2.300	2.117	2.004	1.799
2	1.748	1.733	1.955	1.967	1.867
3	1.901	1.772	1.920	2.056	1.694
4	1.670	1.536	1.655	1.679	1.666
5	1.714	1.592	1.796	1.740	1.682
6	1.707	1.542	1.636	1.489	1.576
7	1.748	1.360	1.706	1.490	1.548
8	1.525	1.284	1.511	1.351	1.361
9	1.630	1.377	1.510	1.306	1.313
H	1.426	1.198	1.279	1.117	1.215
L-H	0.750	1.102	0.839	0.887	0.584
t-stat	**6.24**	**8.87**	**7.23**	**7.99**	**5.57**

Panel B: Profit-Firm sample

Decile	*Accruals*	*ΔNOA*	*NOA*	*CFO*	*FCF*
L	1.318	1.428	1.417	1.293	1.126
2	1.360	1.356	1.328	1.520	1.191
3	1.378	1.354	1.397	1.373	1.415
4	1.348	1.228	1.312	1.540	1.277
5	1.354	1.306	1.327	1.308	1.240
6	1.224	1.139	1.348	1.242	1.237
7	1.276	1.067	1.252	1.209	1.242
8	1.228	1.030	1.143	1.088	1.193
9	1.228	1.011	1.198	1.006	1.086
H	1.049	0.890	1.004	0.873	0.973
L-H	0.269	0.538	0.413	0.421	0.153
t-stat	**3.46**	**6.89**	**4.60**	**5.82**	**2.42**

All variables are defined in Table 1. Every month between July 1964 and December 2002, firms are assigned into decile portfolios based on the magnitude of *Accruals*, *ΔNOA*, *NOA*, *CFO*, or *FCF* from the nearest available fiscal year. To allow for a minimum of a four-month lag between the fiscal year-end and the return month, all returns are measured from the 5th month to the 16th month after the fiscal year-end. The hedge portfolio consists of a long position in the highest ranked portfolio and an offsetting short position in the lowest ranked portfolio. The time-series average of return skewness for each decile is reported for the full sample (Panel A) and for the profit firm sample (Panel B). Bolded t-statistics indicate significance at a less than 5 percent level (two-tailed).

LTS 1 percent trimming deleted 2,890 positive monthly abnormal returns that sum to 292,598 percent but only eliminated 499 negative observations that sum to −31,941%. In comparison, the highest decile loses 1,738 positive observations totaling 153,792 percent, but only 265 negative observations totaling −17,472 percent. Decile 5, on the other hand, has magnitudes that are less than those of decile H. The net effect is that the lowest decile suffers the greatest truncation bias (260,657 percent), followed by the highest decile of (136,320 percent), whereas the middle decile 5 suffers the least (80,154 percent). In relative magnitudes, the truncation bias of the lowest decile is

TABLE 3
Truncation Bias in Full Sample and Profit Sample

Panel A: ***Accruals***

	All-Firm Sample					Profit-Firm Sample				
	Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted			Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted		
Decile	**#**	**Sum Ret.**	**#**	**Sum Ret.**	**Net Returns**	**#**	**Sum Ret.**	**#**	**Sum Ret.**	**Net Returns**
1	2,890	292,598	499	−31,941	260,657	1,431	96,128	283	−14,152	81,976
2	1,713	161,118	215	−13,794	147,324	926	64,462	191	−9,609	54,853
3	1,262	125,573	153	−9,699	115,874	860	58,192	126	−5,882	52,310
4	1,047	96,077	101	−6,509	89,568	725	48,233	128	−6,308	41,925
5	944	87,094	109	−6,940	80,154	759	49,832	146	−7,089	42,743
6	995	90,153	92	−5,906	84,247	770	49,528	147	−7,210	42,318
7	1,104	102,492	131	−8,545	93,947	917	60,691	174	−8,749	51,942
8	1,099	99,267	126	−8,163	91,104	1,039	69,555	223	−10,873	58,682
1	1,362	124,961	129	−8,074	116,887	1,248	83,979	274	−13,351	70,628
10	1,738	153,792	265	−17,472	136,320	1,689	109,573	472	−23,701	85,872
(1 − 10)	1,152	138,806	234	−14,469	124,337	−258	−13,444	−189	9,548	−3,896

Panel B: **Δ*NOA***

	All-Firm Sample					Profit-Firm Sample				
	Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted			Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted		
Decile	**#**	**Sum Ret.**	**#**	**Sum Ret.**	**Net Returns**	**#**	**Sum Ret.**	**#**	**Sum Ret.**	**Net Returns**
1	2,723	283,196	323	−21,179	262,017	1,381	96,718	214	−10,688	86,029
2	1,557	150,617	175	−11,501	139,116	1,034	69,902	168	−8,215	61,686
3	1,151	112,271	119	−7,724	104,547	754	50,951	121	−5,896	45,055
4	875	78,993	87	−5,489	73,504	724	47,447	113	−5,636	41,811

(continued on next page)

TABLE 3 (continued)

	All-Firm Sample					Profit-Firm Sample				
	Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted			Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted		
Decile	#	Sum Ret.	#	Sum Ret.	Net Returns	#	Sum Ret.	#	Sum Ret.	Net Returns
5	809	72,976	83	−5,307	67,669	699	49,605	124	−6,317	43,288
6	826	78,730	85	−5,589	73,140	747	50,900	145	−7,083	43,816
7	737	66,862	85	−5,560	61,301	705	46,130	133	−6,114	40,016
8	906	77,423	111	−7,201	70,223	837	54,398	198	−9,521	44,877
1	1,046	93,151	144	−9,214	83,937	970	62,259	227	−11,130	51,129
10	1,314	112,835	239	−15,362	97,473	1,157	73,539	376	−18,551	54,988
(1 − 10)	1,409	170,361	84	−5,817	164,544	224	23,178	−162	7,863	31,041

Panel C: ***NOA***

	All-Firm Sample					Profit-Firm Sample				
	Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted			Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted		
Decile	#	Sum Ret.	#	Sum Ret.	Net Returns	#	Sum Ret.	#	Sum Ret.	Net Returns
1	2,756	275,237	343	−22,289	252,948	1,414	96,572	277	−13,191	83,381
2	2,138	206,631	256	−16,233	190,398	1,123	74,165	206	−10,095	64,070
3	1,549	146,018	201	−12,886	133,132	1,069	71,500	158	−7,561	63,939
4	1,243	113,616	100	−6,561	107,055	908	61,373	168	−8,156	53,217
5	1,065	104,328	135	−8,639	95,689	892	58,934	149	−7,286	51,648
6	953	85,944	109	−6,974	78,970	879	60,300	164	−7,650	52,650
7	895	84,866	98	−6,134	78,732	874	58,659	161	−8,186	50,473
8	925	84,696	114	−7,325	77,371	882	58,216	200	−9,915	48,301
1	1,102	98,478	158	−10,400	88,078	1,022	66,176	224	−11,363	54,813
10	1,533	133,286	301	−19,211	114,075	1,294	83,935	464	−23,704	60,231
(1 − 10)	1,223	141,951	42	−3,078	138,873	120	12,637	−187	10,513	23,150

(continued on next page)

TABLE 3 (continued)

Panel D: *CFO*

	All-Firm Sample					Profit-Firm Sample				
	Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted			Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted		
Decile	#	Sum Ret.	#	Sum Ret.	Net Returns	#	Sum Ret.	#	Sum Ret.	Net Returns
1	3,841	374,879	557	−36,310	338,569	1,988	135,285	471	−24,095	111,190
2	2,787	272,123	338	−22,074	250,049	1,429	100,556	297	−14,834	85,722
3	1,893	183,016	232	−14,786	168,230	1,111	73,077	229	−11,127	61,950
4	1,359	125,596	149	−9,554	116,042	925	64,492	204	−10,534	53,958
5	935	86,835	116	−7,478	79,357	869	57,631	144	−6,756	50,875
6	750	67,918	94	−6,158	61,760	774	51,138	132	−6,746	44,392
7	678	59,198	73	−4,583	54,615	762	49,994	125	−6,376	43,618
8	593	51,066	61	−3,934	47,132	728	46,755	134	−6,229	40,526
1	601	52,387	60	−3,799	48,588	763	48,086	148	−7,159	40,927
10	747	62,067	110	−6,722	55,345	1,019	63,397	276	−12,895	50,502
(1 − 10)	3,094	312,812	447	−29,588	283,224	969	71,888	195	−11,200	60,688

Panel E: *FCF*

	All-Firm Sample					Profit-Firm Sample				
	Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted			Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted		
Decile	#	Sum Ret.	#	Sum Ret.	Net Returns	#	Sum Ret.	#	Sum Ret.	Net Returns
1	2,723	283,196	323	−21,179	262,017	1,381	96,718	214	−10,688	86,029
2	1,557	150,617	175	−11,501	139,116	1,034	69,902	168	−8,215	61,686
3	1,151	112,271	119	−7,724	104,547	754	50,951	121	−5,896	45,055
4	875	78,993	87	−5,489	73,504	724	47,447	113	−5,636	41,811
5	809	72,976	83	−5,307	67,669	699	49,605	124	−6,317	43,288
6	826	78,730	85	−5,589	73,140	747	50,900	145	−7,083	43,816

(continued on next page)

TABLE 3 (continued)

	All-Firm Sample					Profit-Firm Sample				
	Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted			Positive Abnormal Return Obs. Deleted		Negative Abnormal Return Obs. Deleted		
Decile	#	Sum Ret.	#	Sum Ret.	Net Returns	#	Sum Ret.	#	Sum Ret.	Net Returns
7	737	66,862	85	−5,560	61,301	705	46,130	133	−6,114	40,016
8	906	77,423	111	−7,201	70,223	837	54,398	198	−9,521	44,877
1	1,046	93,151	144	−9,214	83,937	970	62,259	227	−11,130	51,129
10	1,314	112,835	239	−15,362	97,473	1,157	73,539	376	−18,551	54,988
(1 − 10)	1,409	170,361	84	−5,817	164,544	224	23,178	−162	7,863	31,041

All variables are as defined in Table 1. Every month between July 1964 to December 2002, firms are assigned into decile portfolios based on the magnitude of *Accruals*, *ΔNOA*, *NOA*, *CFO*, or *FCF* available from the nearest fiscal year. To allow for a minimum of a four-month lag between the fiscal year-end and the return month, all returns are measured from the 5th month to the 16th month after the fiscal year-end. The hedge portfolio consists of a long position in the highest ranked portfolio and an offsetting short position in the lowest ranked portfolio. The total number of observations and sum of abnormal returns deleted (expressed in percentage) are reported separately for positive and negative abnormal return observations. The hedge portfolio returns (1–10) differ slightly from the amount of the difference between portfolios 1 and 10 because of a rounding error.

3.25 times larger than the middle decile, and is almost double that of the highest decile. These results indicate that the downward truncation bias is an inverted U-shape that is asymmetric and particularly strong in the lower deciles. Similar effects are noticed for the other accounting anomalies. We next use regression analysis to examine whether the truncation bias from trimming dominates the original relation to drive this inverted U-shape pattern.

Pooled OLS Regression Tests of Truncation Bias on Accounting Anomalies

Table 4 presents the regression results of returns on risk controls and predictor deciles for the five anomalies. Each panel contains three regression results for each of the anomalies. The first is the pooled cross-sectional and time-series OLS regression of returns on the predictor deciles after LTS 1 percent trimming and *Size* is the only control variable for risk to facilitate comparison with KLW. The second and third regressions use the Fama-Macbeth procedure and a full set of risk controls that are now standard in the asset-pricing literature. We discuss the pooled OLS regression for the trimmed sample in Table 4 in this subsection.

For each of the accounting predictor (*Accruals*, *ΔNOA*, *NOA*, *CFO*, and *FCF*) regressions, the included test variables are the set of indicator variables *DL*, *D2*, *D3*, *D4*, *D7*, *D8*, *D9*, and *DH*. *DL* is 1 if the firm is in the lowest predictor decile in a particular calendar month, and 0 otherwise. *DH* is the indicator variable for the highest predictor decile, and *D#*, where # ranges from 2 to 9, is the indicator variable for the corresponding # decile. *D5* and *D6* are excluded, such that the estimated coefficients on the included test portfolios measure the incremental abnormal returns from the base portfolio containing *D5* and *D6*.

Consistent with KLW, the pooled OLS regressions in the trimmed sample exhibit an inverted-U shape for the *Accruals*, *ΔNOA*, and *NOA* anomalies. *DL* and *DH* coefficients are statistically significant and negative, and are more negative than other decile coefficients. The negative coefficients for *DL* and other lower deciles suggest that the downward truncation bias from trimming is large enough to swamp the underlying relation between accruals-related predictors and future returns.

On the other hand, the results for *CFO* and *FCF* suggest that the downward truncation bias amplifies the negative returns for the lower deciles and has a relatively smaller negative effect on the upper deciles. The relation between *CFO* and returns is monotonic up to decile 8, and deciles 9 and 10 coefficients, although lower than decile 8, remain positive and significant. The relation between *FCF* and returns remains positive and monotonic throughout, and therefore trimming does not lead to a rejection of the behavioral theory for *FCF*.

Fama-Macbeth Regression and Additional Asset-Pricing Controls

We discuss the results for the next two regressions in each of the panels in Table 4 in this subsection. In the finance literature, the Fama-Macbeth regression method is commonly used to investigate cross-sectional return predictabilities (see Campbell et al. 1997; Daniel and Titman 1997). In a typical asset-pricing application of the Fama-Macbeth method, researchers regress monthly stock returns on annual values of lagged firm-level financial variables. This procedure has two advantages over the simple pooled OLS regression. Fama-Macbeth standard errors are more efficient than the OLS estimates (Petersen 2009; Gow et al. 2010),[10] whereas OLS pooled regression test statistics are biased because of cross-sectional correlation of residuals. Especially

---

[10] A few authors suggest adjusting the Fama-Macbeth standard error for the estimated first order auto-correlation of the estimated slope coefficients to account for any possible firm effect (Newey-West adjusted standard error). Petersen (2009) shows that the adjustment may not work under certain circumstances. Also see Gow et al. (2010).

TABLE 4

Fama-Macbeth and OLS Regressions of Returns on Asset-Pricing Controls and Decile Rankings of Accounting Predictors before and after 1 Percent LTS Trimming (t-statistics in italics)

Panel A: ***Accruals***

Intercept	**Ln(*Size*)**	**Ln(*B/M*)**	**Ret(−1: −1)**	**Ret(−12: −1)**	**Ret(−36: −13)**	***DL***	***D2***	***D3***	***D4***	***D7***	***D8***	***D9***	***DH***
Ordinary Least Squares-LTS (Least-Trimmed Square) with *Size* as the only Control													
0.0064	0.0001					−0.0115	−0.0027	0.0006	0.0004	−0.0012	−0.0023	−0.0056	−0.0138
8.81	*1.46*					***−24.37***	***−5.80***	*1.33*	*0.85*	***−2.52***	***−4.84***	***−11.92***	***−29.40***
Fama-Macbeth													
0.0296	−0.0014	0.0029	−0.0697	0.0055	−0.0018	0.0001	0.0004	0.0020	0.0008	−0.0005	−0.0013	−0.0031	−0.0082
4.46	***−2.89***	***4.18***	***−16.35***	***3.33***	***−2.95***	*0.09*	*0.53*	***3.70***	*1.78*	*−0.93*	***−2.14***	***−4.40***	***−8.17***
Fama-Macbeth-LTS													
−0.0082	0.0015	0.0050	−0.0632	0.0104	0.0002	−0.0031	0.0003	0.0016	0.0011	0.0001	0.0003	−0.0017	−0.0070
−1.36	***3.58***	***7.99***	***−16.28***	***6.40***	*0.41*	***−3.06***	*0.54*	***3.39***	***2.59***	*0.16*	*0.47*	***−2.58***	***−7.27***

Panel B: ***ΔNOA***

Intercept	**Ln(*Size*)**	**Ln(*B/M*)**	**Ret(−1: −1)**	**Ret(−12: −1)**	**Ret(−36: −13)**	***DL***	***D2***	***D3***	***D4***	***D7***	***D8***	***D9***	***DH***
Ordinary Least Square-LTS (Least-Trimmed Square) with *Size* as the only Control													
0.0008	0.0005					−0.0023	0.0045	0.0060	0.0060	0.0036	0.0011	−0.0017	−0.0099
2.24	***8.50***					***−4.99***	***9.74***	***13.15***	***13.01***	***7.94***	***2.39***	***−3.62***	***−21.6***
Fama-Macbeth													
0.01863	−0.0012	0.0027	−0.0675	0.0057	−0.0015	0.0050	0.0033	0.0032	0.0017	−0.0004	−0.0018	−0.0022	−0.0082
5.02	***−2.63***	***3.88***	***−15.70***	***3.54***	***−2.56***	***5.53***	***5.11***	***5.71***	***3.22***	*−0.75*	***−3.37***	***−3.72***	***−10.38***
Fama-Macbeth-LTS													
0.0014	0.0016	0.0048	−0.0628	0.0104	0.0004	0.0013	0.0032	0.0035	0.0022	0.0009	−0.0002	−0.0011	−0.0068
0.39	***3.81***	***7.73***	***−16.16***	***6.61***	*0.67*	*1.69*	***5.99***	***7.22***	***4.73***	***2.00***	*−0.49*	***−1.96***	***−9.25***

(continued on next page)

TABLE 4 (continued)

Panel C: *NOA*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1: −1)	Ret(−12: −1)	Ret(−36: −13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
Ordinary Least Square-LTS (Least-Trimmed Square) with *Size* as the Only Control													
0.0030	0.0004					−0.0055	−0.0037	−0.0015	−0.0005	−0.0008	−0.0037	−0.0064	−0.0170
4.28	***7.29***					***−11.67***	***−7.92***	***−3.25***	*−1.10*	*−1.72*	***−7.92***	***−13.79***	***−36.41***
Fama-Macbeth													
0.0261	−0.0010	0.0036	−0.0701	0.0053	−0.0017	0.0055	0.0029	0.0015	0.0011	−0.0009	−0.0035	−0.0038	−0.0106
3.89	***−2.19***	***5.44***	***−16.48***	***3.19***	***−2.85***	***4.62***	***3.24***	***2.40***	***2.03***	*−1.80*	***−7.03***	***−6.28***	***−11.12***
Fama-Macbeth-LTS													
−0.0106	0.0018	0.0056	−0.0634	0.0103	0.0003	0.0024	0.0011	0.0007	0.0005	−0.0006	−0.0027	−0.0034	−0.0102
−1.73	***4.23***	***9.24***	***−16.48***	***6.33***	*0.53*	***2.33***	*1.32*	*1.25*	*1.11*	*−1.50*	***−6.34***	***−6.32***	***−11.21***

Panel D: *CFO*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1: −1)	Ret(−12: −1)	Ret(−36: −13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
Ordinary Least Square-LTS (Least-Trimmed Square) with *Size* as the Only Control													
0.0171	−0.0008					−0.0265	−0.0156	−0.0084	−0.0042	0.0022	0.0039	0.0032	0.0021
23.38	***−13.65***					***−56.13***	***−32.91***	***−17.86***	***−8.97***	***4.68***	***8.28***	***6.89***	***4.51***
Fama-Macbeth													
0.0315	−0.0016	0.0036	−0.0707	0.0049	−0.0026	−0.0053	−0.0039	−0.0026	−0.0020	0.0017	0.0032	0.0038	0.0055
5.11	***−3.67***	***5.46***	***−16.84***	***3.04***	***−4.29***	***−3.30***	***−3.16***	***−3.04***	***−3.40***	***3.71***	***6.78***	***6.82***	***8.21***
Fama-Macbeth-LTS													
−0.0018	0.0010	0.0051	−0.0650	0.0095	−0.0006	−0.0129	−0.0075	−0.0038	−0.0023	0.0016	0.0030	0.0035	0.0052
−0.32	***2.54***	***8.39***	***−16.86***	***5.94***	*−1.03*	***−8.99***	***−7.10***	***−4.85***	***−4.72***	***3.87***	***6.83***	***7.05***	***8.41***

(continued on next page)

TABLE 4 (continued)

Panel E: ***FCF***

Intercept	**Ln(*Size*)**	**Ln(*B/M*)**	**Ret(−1: −1)**	**Ret(−12: −1)**	**Ret(−36: −13)**	***DL***	***D2***	***D3***	***D4***	***D7***	***D8***	***D9***	***DH***
Ordinary Least Square-LTS (Least-Trimmed Square) with *Size* as the only Control													
0.0078	−0.0003					−0.0252	−0.0137	−0.0072	−0.0029	0.0038	0.0044	0.0056	0.0056
21.93	***−5.92***					***−54.55***	***−29.77***	***−15.61***	***−6.36***	***8.16***	***9.66***	***12.14***	***12.26***
Fama-Macbeth													
0.0177	−0.0018	0.0055	−0.0682	0.0049	−0.0025	−0.0095	−0.0048	−0.0034	−0.0025	0.0017	0.0024	0.0033	0.0051
4.12	***−4.04***	***3.75***	***−15.99***	***3.11***	***−4.23***	***−6.28***	***−4.22***	***−4.52***	***−4.57***	***3.28***	***4.63***	***5.56***	***7.66***
Fama-Macbeth-LTS													
−0.0031	0.0010	0.0101	−0.0634	0.0094	−0.0005	−0.0134	−0.0064	−0.0031	−0.0015	0.0027	0.0032	0.0045	0.0061
−0.78	***2.64***	***7.68***	***−16.57***	***6.08***	*−0.92*	***−9.67***	***−6.55***	***−4.61***	***−3.05***	***5.71***	***7.03***	***8.40***	***10.01***

Every month *t* between July 1964 and December 2002, the cross-section of stock returns are regressed on Ln(*Size*) where *Size* is defined as the log of the firm's market capitalization; Ln(*B/M*), which is the log of the book-to-market ratio; the previous month's return on the stock, denoted Ret(−1:−1); the previous year's return on the stock from month *t*−12 to *t*−2, denoted Ret(−12:−2); the return on the stock starting from month *t*−36 to *t*−13, denoted Ret(−36:−13) and the decile rank of the financial variable of interest. The return measurement period begins four months after the firm's fiscal year-end. The indicator variable *DL* is set to 1 if the firm is in the lowest predictor decile in month *t*, 0 otherwise. The indicator variable *DH* is set to 1 if the firm is in the highest predictor decile in month *t*, 0 otherwise. *D2* to *D9* are indicator variables set to 1 if the firm is in the corresponding accounting predictor decile. For the Ordinary Least Squares-Least Trimmed Square (LTS) regression, we truncate the observations whose square of residuals from the Ordinary Least Squares regressions are at the top 1 percent level, then rerun the Ordinary Least Squares regression. For the Fama-Macbeth regression, the time-series average of the monthly coefficient estimates is reported. For the Fama-Macbeth-Least Trimmed Square (LTS) regression, every month, we truncate observations whose square of residuals from the original Fama-Macbeth regression are at the top 1 percent level, and then we rerun the Fama-Macbeth regression. Bolded t-statistics indicate significance at less than 5 percent level (two-tailed).

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

relevant for the current application, the Fama-Macbeth procedure also reduces the likelihood that an extreme monthly return would exert an undue influence on the regression estimates because the coefficients are averaged over all monthly regressions.[11]

Monthly cross-sectional regressions of returns on the decile ranks of the predictor variables (*Accruals*, *ΔNOA*, *NOA*, *CFO*, and *FCF*) are run for the untrimmed sample in the second regression and for the trimmed sample in the third regression. The time-series-averaged coefficients and their t-statistics are reported in Table 4. Following Hirshleifer et al. (2004), we include a full set of standard asset-pricing controls in addition to size, such as book-to-market, past one-month return (for the Jegadeesh [1990] contrarian effect), past one-year return excluding latest month (for the Jegadeesh and Titman [1993] momentum effect), and past three-year returns excluding the latest year (using returns from months −36 to −13 for the DeBondt and Thaler [1985] long-term reversal effect).[12,13]

The empirical results in the second and third regressions using the Fama-Macbeth method differ from those in the first regression using pooled OLS regression in Table 4. The t-statistics are much smaller for the Fama-Macbeth regressions than for the pooled OLS regressions, which are likely overstated because of cross-sectional correlation of the regression residuals.[14] The Fama-Macbeth regressions for the full sample confirm the presence of the five anomaly relations documented in the prior literature. The accruals-related anomaly relations with future returns are negative, while the cash-flow-related anomalies are positive. We focus on the comparison of the results with the trimmed sample next.

The Fama-Macbeth regression method and inclusion of additional standard asset-pricing controls weaken the inverted-U relation for *Accruals* in the trimmed sample relative to the pooled OLS trimmed sample. Only the *Accruals* anomaly exhibits a similar inverted U-shape after trimming. For the *ΔNOA* anomaly, the lowest decile has positive incremental abnormal returns that are insignificantly different from the middle decile, and the abnormal returns from decile 3 onwards increase monotonically. The *NOA* anomaly is even more robust, maintaining positive and significantly larger abnormal returns relative to the middle deciles, and monotonically increasing abnormal returns after trimming. The greater robustness of the *NOA* anomaly with respect to trimming relative to the *Accruals* anomaly is consistent with previous finding that the *NOA* anomaly is stronger than the operating accrual anomaly, especially on the long side (Hirshleifer et al. 2004).

---

[11] KLW suggest that their results are robust with respect to using logged buy and hold annual returns (BHAR). Taking logarithms shrinks extreme large positive numbers, and exaggerates large negative numbers. Therefore, the effect of using log BHAR is similar to the effect obtained using LTS trimming. Furthermore, the simple return on a portfolio of assets is the simple weighted average of the individual asset returns. Logged returns do not share this convenient property since the log of a sum is not the same as the sum of logs (Campbell et al. 1997). When logged returns are used, the regression coefficients do not have a clear interpretation as portfolio returns, and therefore logged returns are not commonly used in the cross-sectional, asset-pricing literature (Campbell et al. 1997, 12).

[12] Campbell and Vuolteenaho (2004) suggest two additional risk factors that represent news about future dividends and news about future discount rate. Hirshleifer et al. (2010) and our untabulated results find that these factors are not significant, and therefore we do not consider them as additional risk controls.

[13] An earlier version of the study showed the incremental effect of changing the regression method from pooled OLS to Fama-Macbeth procedure separately first, before adding a fuller set of risk controls, beyond *Size*. The switch in regression method to Fama Macbeth reduces the magnitude of the truncation bias from pooled OLS so that the coefficient for *DL* is halved, although still statistically significant for *Accruals*, and is not significant for *NOA*.

[14] Fama-Macbeth t-statistics also tend to have lower power because they are calculated using a much smaller number of observations relative to the pooled OLS regression t-statistics. There are only 462 monthly cross-sectional regression coefficients in our sample period, which is much fewer than the number of total observations in the pooled OLS regression (about 1.6 million).

In consequence, the truncation bias is not strong enough to dominate the *NOA* anomaly in any range of the deciles.

For the *CFO* and *FCF* anomalies, trimming has little effect on the shape of the relation. Both anomalies continue to exhibit strict monotonicity across deciles, and the size of the anomalies are actually larger after trimming. The trading profit for *CFO* increases from 1.08 percent per month in the untrimmed sample to 1.81 percent in the trimmed sample, and that for *FCF* increases from 1.95 percent to 3.08 percent. The evidence confirms the effect illustrated previously in Figure 2 that trimming strengthens the cash-flow-based anomalies and does not eliminate monotonicity of the relation.

In sum, the results from all five anomalies suggest that appropriate regression methods and risk controls are important to infer presence of anomalies. The monotonic pattern in all the anomalies, except the operating accrual anomaly, is robust to trimming using monthly returns, using the Fama-Macbeth regression method to obtain better behaved t-statistics, and incorporating a more comprehensive set of risk controls.

IV. ALTERNATIVE TRIMMING METHODS

In this section, we consider two alternative sample selection methods to reduce the variation in returns skewness across predictor deciles, thereby reducing, although not eliminating, the truncation bias effects from trimming. We are not suggesting that trimming is appropriate for these subsamples. Rather, the goal of the exercise is to provide counter-examples where monotonicity of anomalies are present in large subsets of the population to offer assurance to researchers concerned that the anomalies may be driven by only a small fraction of the observations. By using alternative methods to implement trimming and yet preserve a large fraction of the population, we provide evidence that KLW's inference that anomalies are not robust to trimming is itself not robust. Table 5 reports the results from a uniform LTS 1 percent trimming at each predictor decile rather than trimming at the overall sample level. Tables 6 and 7 present the analysis for when loss firms are pruned from the sample.

Uniform LTS Trimming across Accounting Predictor Deciles

In this procedure, we first run the monthly cross-sectional regressions for the entire sample, then rank the squared residuals in each accounting predictor decile, delete the top 1 percent observations within each decile, and re-run the regression with the remaining 99 percent of the sample. This procedure gives a higher chance of preserving the dominance of the underlying returns relation across predictor deciles, yet still eliminates the influence of extreme observations within each decile.

Table 5 reports the results. For the *Accruals* sample, a uniform LTS trimming of the top 1 percent across each *Accruals* decile leaves a near monotonic negative relation between *Accruals* and abnormal returns that is similar to the full-sample second regression in Table 4, Panel A as well as the full sample pooled OLS regression without trimming (not reported in Table 4 for brevity).

The results of uniform trimming across deciles in the remaining panels of Table 5 for the other four anomalies are also similar to the second regression in the corresponding panels of Table 4 for the full sample without trimming. The result that the relation between predictors and returns remains monotonic in the uniformly trimmed sample comprising of 99 percent of the population illustrates the robustness of behavioral explanations for anomalies, consistent with Core's (2006, 346) conjecture that "[h]ad KLW found an even distribution of extreme returns across the deciles, they would have found that the remaining data support the earnings fixation hypothesis, and would have concluded that the evidence for the hypothesis was robust."

TABLE 5

Sample after 1 Percent LTS Trimming in Each Decile
Fama-Macbeth Regressions of Returns on Asset-Pricing Controls and Decile Ranking of Accounting Predictors
(t-statistics in italics)

Panel A: *Accruals*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1:−1)	Ret(−12:−1)	Ret(−36:−13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
−0.0021	0.0009	0.0095	−0.0608	0.0100	−0.0004	0.0011	0.0014	0.0020	0.0012	0.0005	0.0005	−0.0014	−0.0057
−0.53	***2.26***	***6.73***	***−14.95***	***6.28***	*−0.72*	*0.94*	***2.00***	***3.90***	***2.69***	*1.08*	*0.83*	***−2.13***	***−5.76***

Panel B: *ΔNOA*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1:−1)	Ret(−12:−1)	Ret(−36:−13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
0.0048	0.0010	0.0042	−0.0616	0.0095	−0.0001	0.0050	0.0037	0.0030	0.0014	−0.0003	−0.0014	−0.0019	−0.0073
1.35	***2.36***	***6.47***	***−15.36***	***6.01***	*−0.11*	***5.79***	***6.26***	***5.77***	***2.84***	*−0.73*	***−2.86***	***−3.39***	***−9.54***

Panel C: *NOA*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1:−1)	Ret(−12:−1)	Ret(−36:−13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
−0.0028	0.0011	0.0106	−0.0611	0.0097	−0.0003	0.0047	0.0030	0.0012	0.0007	−0.0007	−0.0029	−0.0033	−0.0095
−0.70	***2.72***	***7.86***	***−15.22***	***6.17***	*−0.57*	***3.90***	***3.29***	***2.00***	*1.42*	*−1.66*	***−6.46***	***−6.13***	***−10.21***

Panel D: *CFO*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1:−1)	Ret(−12:−1)	Ret(−36:−13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
−0.0001	0.0005	0.0100	−0.0622	0.0089	−0.0013	−0.0062	−0.0036	−0.0021	−0.0012	0.0012	0.0026	0.0034	0.0045
−0.02	*1.32*	***7.34***	***−15.74***	***5.74***	***−2.27***	***−3.55***	***−2.79***	***−2.44***	***−2.14***	***2.82***	***6.01***	***7.09***	***7.59***

(continued on next page)

TABLE 5 (continued)

Panel E: *FCF*

Intercept	Ln(Size)	Ln(B/M)	Ret(−1:−1)	Ret(−12:−1)	Ret(−36:−13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
0.0015	0.0005	0.0084	−0.0625	0.0085	−0.0011	−0.0093	−0.0045	−0.0030	−0.0022	0.0012	0.0021	0.0032	0.0049
0.38	*1.20*	***6.16***	***−15.64***	***5.47***	*−1.90*	***−6.15***	***−4.09***	***−4.03***	***−4.28***	***2.47***	***4.35***	***5.69***	***7.83***

Every month *t* between July 1964 and December 2002, the cross-section of stock returns are regressed on Ln(*Size*) where *Size* is defined as the log of the firm's market capitalization; Ln(*B/M*), which is the log of the book-to-market ratio; the previous month's return on the stock, denoted Ret(−1:−1); the previous year's return on the stock from month *t*−12 to *t*−2, denoted Ret(−12:−2); the return on the stock starting from month *t*−36 to *t*−13, denoted Ret(−36:−13), and the decile rank of the financial variable of interest. There is a minimum of a four-month gap between the fiscal year-end and the return month in month *t* regressions. Every month, we truncate observations whose square of residuals from the Fama-Macbeth regression are at the top 1 percent level within each predictor decile; and then we rerun the Fama-Macbeth regression. The time-series average of the monthly coefficient estimates is reported. Bolded t-statistics indicate significance at less than 5 percent level (two-tailed).

TABLE 6

Distribution of Returns between Profit versus Loss Sample and the Frequency of Losses across Accounting Predictor Deciles

Panel A: Distribution of Monthly Returns (One-Year-Ahead) for Profit versus Loss Firms

Min	1%	5%	10%	25%	Median	75%	90%	95%	99%	Max
Profit Firms										
−98.10%	−33.30%	−19.60%	−13.90%	−6.30%	0.00%	7.40%	16.90%	25.00%	48.70%	1400.00%
	Mean		Variance		Skewness		Kurtosis			
	1.28%		0.023		3.39		119			
Loss Firms										
−97.10%	−52.00%	−33.00%	−25.00%	−13.30%	−1.60%	10.00%	28.00%	44.80%	100%	1900.00%
	Mean		Variance		Skewness		Kurtosis			
	0.91%		0.085		5.84		148			

Panel B: Frequency of Loss Firms across Accounting Predictor Deciles

Decile Ranking	*Accruals*	*ΔNOA*	*NOA*	*CFO*	*FCF*
1	42.2%	43.8%	39.6%	58.6%	49.6%
2	25.1%	25.3%	28.5%	43.1%	38.0%
3	17.7%	17.1%	20.6%	29.4%	26.7%
4	13.9%	13.5%	16.2%	17.0%	18.8%
5	12.4%	10.1%	13.2%	9.3%	12.3%
6	12.3%	9.6%	11.0%	5.4%	7.7%
7	13.0%	9.5%	9.9%	3.8%	5.0%
8	12.4%	10.1%	9.7%	3.2%	4.2%
9	12.3%	11.0%	10.4%	2.7%	3.7%
10	14.0%	15.9%	16.3%	2.9%	3.6%

Panel A reports the return distribution statistics for the profit versus the loss firm sample. The monthly returns are measured from the 5th to the 16th month after the fiscal year-end. Panel B reports the frequency of loss firms across *Accruals*, *ΔNOA*, *NOA*, *CFO*, and *FCF* deciles. Every month, the number of loss firms to the number of total firms is calculated for every decile; then the time-series average of the frequency of loss firms is reported.

Truncation Bias from Loss Firms

Recent literature suggests that loss firms differ markedly from profit firms, including their return distributions.[15] The prevalence of losses has risen markedly, from about 3 percent of firms in the early 1960s to over 40 percent in 2000 among firms covered by Compustat. The economic circumstances of loss firms suggest high expected positive skewness for returns, both because loss firms are likely to have high earnings volatility (and raw stock returns are bounded below by −100 percent), and because such firms may in many cases be like lottery gambles in which there is a long upper tail of possible favorable outcomes.[16]

---

[15] Klein and Marquardt (2006) report evidence that the percentage of losses for U.S. firms is significantly related to accounting conservatism, Compustat coverage of small firms, real firm performance, and business cycle factors. Several authors document that the stock price behavior of loss firms differs significantly from that of profit firms. There is no significant contemporaneous return-earnings relation for losses (Hayn 1995), post-earnings announcement drift is weaker among loss than among profit firms (Narayanamoorthy 2006), and the *Accruals* anomaly disappears in a loss sample (Dopuch et al. 2010; Zhang and Shi 2007). KLW observe that most firms with extreme *positive* BHARs experience poor financial performance (losses). In contrast, Beneish et al. (2001) find that an indicator variable for losses over the past four quarters *negatively* forecast returns only in a sample trimmed of extreme performing firms.

Table 6, Panel A confirms the difference between monthly return distributions of loss versus profit firms. The median return for loss firms (−1.60 percent) is lower than for profit firms (0%). The return distribution of loss firms deviates from normality far more than that of profit firms. The variance is 3.7 times larger and skewness 1.7 times higher for loss firms than for profit firms. The returns spread between profit and loss firms at the 1st percentile is about 20 percent, whereas at the 99th percentile is about 50 percent. In sum, loss firms are more right-skewed than are profit firms.

Table 6, Panel B reports the variation in the frequency of loss firms across predictor deciles for all five anomalies. The decrease in the frequency of losses across the predictor deciles is dramatic, is an inverted U-shape for the accruals-related predictors, and is monotonically declining for the cash-flow-related predictors. The average frequency of losses in the bottom *Accruals*, *ΔNOA*, and *NOA*, decile portfolios respectively is 42 percent, 44 percent, 40 percent, decreasing to only 12 percent, 10 percent, and 11 percent for decile 6. In contrast, the frequency of losses remains fairly constant and low (about 10–16 percent) across the five highest deciles. For the *CFO* and *FCF* deciles, the frequency of losses declines from 59 percent and 50 percent, respectively, in the bottom decile to only 3 percent and 4 percent, respectively, in the top decile.

The variations in loss firms across predictor deciles mimic the return skewness patterns in Table 2 and the truncation bias patterns in Table 3. Returning to Panel B of Table 2, the differences in skewness between the highest and the lowest *Accruals*, *ΔNOA*, *NOA*, *CFO*, and *FCF* deciles are reduced by more than half from 0.75, 1.10, 0.94, 0.85, and 0.58 in the full sample to 0.26, 0.54, 0.41, 0.42, and 0.15 in the profit sample, respectively.

The magnitude of the truncation bias also shrinks dramatically when loss firms are eliminated. Comparing the truncation bias across deciles for the general sample on the left-hand side of the table versus the profit sample on the right-hand side in Table 3, the truncation bias is much less severe for the profit sample. Trimming induces a huge drop in total hedge returns in the full sample by 124,337 percent, but actually slightly increases the total abnormal returns between top and bottom deciles in the profit sample in relative terms by −3,896 percent. For the other predictors, the difference in downward truncation bias from trimming the bottom and top deciles is about five to six times as large for the full sample relative to the profit sample.[17]

Overall, the smaller right-skewness leading to smaller downward truncation bias, as well as the smaller variation in downward truncation bias across deciles suggest that anomalies may be more robust to trimming in the profit sample than in the full sample. Before we examine whether this is the case, we consider how trimming can induce a spurious predictability for returns by loss firms to illustrate the caution expressed by Kothari et al. (2005).

---

[16] Knez and Ready (1997) proposed the "Turtle Egg" hypothesis to explain why extreme large return observations are concentrated in small firms. The name comes from the fact that mother sea turtles lay many eggs, but few hatch and even fewer hatched turtles make it to the ocean, but those that succeed can grow to be quite large. In the same way, loss firms often do not survive, but some that do may ultimately become large when they are able to turn the firm around.

[17] As previously noted, small firms and highly volatile firms also have high right-skewness of returns. Trimming of large positive returns will unduly affect deciles where they are heavily represented. Untabulated results show that small and highly volatile firms are distributed somewhat evenly in the top and bottom accruals-related predictor deciles unlike loss firms that dominate the bottom deciles.
Evidence of asymmetric downward truncation bias (e.g., two times larger for bottom than for top *Accruals* decile) in Table 3 for LTS trimming in the full sample suggests that loss firms are more likely the source for the asymmetric truncation bias rather than small or highly volatile firms. Therefore, we consider trimming in a sample with loss firms removed to illustrate the falseness of KLW's inference that the accruals-related anomalies are driven by just 1 percent of the sample.

TABLE 7

Profit Firms Only Sample
Fama-Macbeth Regressions of Returns on Asset-Pricing Controls and Decile Ranking of Accounting Predictors after LTS 1 Percent Trimming
(t-statistics in italics)

Panel A: *Accruals*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1: −1)	Ret(−12: −1)	Ret(−36: −13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
−0.0013	0.0009	0.0031	−0.0603	0.0101	−0.0003	0.0012	0.0015	0.0014	0.0013	0.0001	−0.0008	−0.0025	−0.0082
−0.24	***2.41***	***5.16***	***−15.77***	***6.39***	*−0.47*	***2.01***	***3.01***	***3.19***	***3.26***	*0.28*	*−1.39*	***−3.71***	***−8.24***

Panel B: Δ*NOA*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1: −1)	Ret(−12: −1)	Ret(−36: −13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
0.0041	0.0010	0.0029	−0.0600	0.0101	−0.0001	0.0032	0.0032	0.0032	0.0018	−0.0002	−0.0007	−0.0017	−0.0064
1.28	***2.64***	***4.93***	***−15.71***	***6.59***	*−0.17*	***5.50***	***6.25***	***7.22***	***3.83***	*−0.40*	*−1.39*	***−3.04***	***−8.80***

Panel C: *NOA*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1: −1)	Ret(−12: −1)	Ret(−36: −13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
−0.0029	0.0011	0.0035	−0.0601	0.0098	−0.0003	0.0032	0.0013	0.0008	0.0006	−0.0009	−0.0026	−0.0034	−0.0087
−0.51	***2.87***	***5.84***	***−15.69***	***6.20***	*−0.53*	***4.82***	***2.47***	***1.98***	*1.57*	***−2.13***	***−5.99***	***−6.36***	***−11.13***

Panel D: *CFO*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1: −1)	Ret(−12: −1)	Ret(−36: −13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
0.0001	0.0009	0.0047	−0.0605	0.0096	−0.0008	−0.0096	−0.0041	−0.0029	−0.0014	0.0015	0.0020	0.0032	0.0039
0.02	***2.36***	***7.09***	***−15.77***	***6.08***	*−1.49*	***−11.87***	***−6.72***	***−6.16***	***−3.22***	***3.68***	***5.24***	***6.40***	***6.38***

(continued on next page)

TABLE 7 (continued)

Panel E: *FCF*

Intercept	Ln(*Size*)	Ln(*B/M*)	Ret(−1: −1)	Ret(−12: −1)	Ret(−36: −13)	*DL*	*D2*	*D3*	*D4*	*D7*	*D8*	*D9*	*DH*
0.0001	0.0008	0.0083	−0.0599	0.0095	−0.0007	−0.0088	−0.0037	−0.0023	−0.0013	0.0014	0.0024	0.0033	0.0047
0.03	***2.22***	***5.93***	***−15.69***	***6.22***	*−1.27*	***−10.96***	***−6.28***	***−4.52***	***−3.06***	***3.21***	***5.55***	***6.77***	***7.54***

This sample consists of profit firms only. Every month *t* between July 1964 and December 2002, the cross-section of stock returns are regressed on Ln(*Size*) where *Size* is defined as the log of the firm's market capitalization; Ln(*B/M*), which is the log of the book-to-market ratio; the previous month's return on the stock, denoted Ret(−1: −1); the previous year's return on the stock from month *t*−12 to *t*−2, denoted Ret(−12: −2); the returns on the stock starting from month *t*−36 to *t*−13, denoted Ret(−36: −13), and decile ranks of accounting predictors. There is a minimum of a four-month gap between the fiscal year-end and the return month in month *t* regressions. Every month, we truncate observations whose square of residuals from the Fama-Macbeth regression are at the top 1 percent level within each predictor decile, and we then rerun the Fama-Macbeth regression. The time-series average of the monthly coefficient estimates is reported. Bolded t-statistics indicate significance at less than 5 percent level (two-tailed).

Spurious Loss Predictability for Returns

To investigate the relation between a loss indicator variable (*Loss* = 1 if the firm experienced a loss, and 0 otherwise) and returns in the subsequent 12 months, we run Fama-Macbeth monthly cross-sectional regressions on *Loss* and the same set of extensive standard risk controls in both trimmed and untrimmed samples. In the untrimmed sample, the *Loss* coefficient is a small −0.23 percent and is not statistically significant. However, *Loss* is a negative and significant predictor of future returns in the trimmed sample; its coefficient is −1.04 percent per month (t = −7.95) and is more than four times larger than in the untrimmed sample, consistent with Beneish et al. (2001). The *Loss* variable continues to be statistically significant even when the anomaly variables (e.g., *Accruals*, *CFO*, and *NOA*) are added to the regression. (The full set of results is not reported in the tables for brevity.)

These results are consistent with our conjecture that loss firms suffer from a stronger negative truncation bias on returns because of its greater return right-skewness. Importantly, the seemingly highly significant "loss effect" in the trimmed sample is a spurious consequence of data truncation, confirming Kothari et al. (2005).

Robustness of Anomalies in Profit Firms after LTS 1 Percent Trimming

Returning to the robustness test of anomalies, we now discuss the results of the Fama-Macbeth regressions of returns on the anomaly predictor deciles and standard risk controls for a sample of profit-only firms after LTS 1 percent trimming in Table 7. Panel A of Table 7 shows that the lowest *Accruals* decile is associated with incremental and statistically significant *positive* abnormal future returns that are similar in magnitude to decile 2. From decile 2, the relation between *Accruals* and returns is strictly monotonic. Thus, consistent with our prediction that in a sample in which the truncation bias is largely reduced, the *Accruals* anomaly is robust to trimming.

For both ΔNOA and *NOA*, the hedge profits remain large (about 1 percent and 1.2 percent, respectively), and the relation is monotonic even with trimming in the profit sample. For the *CFO* and *FCF* anomalies, the relation between predictors and returns remain strictly monotonic after trimming. The hedge profits are large and statistically significant (1.35 percent per month for both anomalies) and closer to the hedge profits of the untrimmed full sample (1.81 percent per month for *CFO* and 1.95 percent per month for *FCF* in Table 4) than those from the trimmed full sample.

In sum, KLW's inverted-U relation for the *Accruals* and *NOA* anomalies when data are trimmed is a consequence of a non-random downward truncation bias in samples with highly skewed returns. The return distribution of loss firms are more skewed than that of profit firms, and low *Accruals* or *NOA* firms have a higher incidence of losses than high *Accruals* or *NOA* firms. The robustness of the *Accruals* and *NOA* anomalies, and additionally the ΔNOA, *CFO*, and *FCF* anomalies that were not considered by KLW, with respect to trimming in the profit sample is evidence that the anomalies are more pervasive than KLW suggest. It is not the case that these anomalies are driven by only 1 percent of the sample. Instead, the anomalies are robust across 79 percent of the firms (profit sample).

V. CONCLUSION

Statistical theory teaches that it is invalid to draw inferences from a sample that is selected based upon the values of the dependent variable. Researchers in many sciences take special pains to avoid *ex post* selection bias. Most researchers argue that outliers should only be removed when they represent data errors. In testing robustness of anomalies, Kraft et al. (2006; KLW) trim 1 percent of extreme observations using the LTS procedure and find that the resulting relation between *Accruals* or *NOA* with abnormal returns is an inverted U-shape, which leads them to reject behavioral explanations for these anomalies.

We show that this inverted-U relation is a consequence of the downward truncation bias of right-skewed return distributions that varies nonlinearly with return predictor deciles, a problem that is exacerbated by inadequate risk controls, and regression test methods that do not control appropriately for cross-sectional correlation of residuals. Going beyond the *Accruals* and *NOA* anomalies investigated by KLW, we also examine the robustness of ΔNOA, *CFO*, and *FCF* anomalies.

We find that low accounting predictor deciles are disproportionately loss firms, which have greater positive skewness than profit firms, and therefore are more likely to suffer from truncation bias. In contrast, in the upper five deciles, there is relatively smaller truncation bias. LTS trimming in a sample that includes loss firms exerts stronger downward truncation bias in the lower deciles than in the upper deciles. The asymmetric pattern in truncation bias alters the original monotonic negative relation into an inverted-U relation for the accruals-based anomalies and strengthens the monotonic positive relation for the cash-flow based anomalies.

We consider two alternative trimming methods that are more likely to have smaller variations in truncation bias across accounting predictors as robustness tests of these anomalies. When the LTS trimming procedure is applied uniformly to 1 percent of the largest squared residuals in each predictor decile, the resulting relation between accounting predictors and abnormal returns for 99 percent of the sample is close to monotonic. Similarly, when the 1 percent LTS trimming is applied to a sample limited to only profit firms, which comprise 79 percent of the population, the hedge profits remain statistically significant and the monotonic relation between accounting predictors and abnormal returns is generally preserved. Finally, we also show that trimming induces a spurious loss anomaly.

Our results, together with the warnings in Kothari et al. (2005) and Core (2006), demonstrate that researchers should not trim *ex post* extreme return observations unless such observations are demonstrably due to data error. Trimming valid extreme return observations biases tests of return predictability and anomalies. The estimated trimmed returns of trading strategies are also not achievable based on information available to investors *ex ante*. Furthermore, since R^2s in return prediction regressions are generally very low, *ex post* removal of even a small fraction of observations can substantially alter the distribution of trading strategy profits. Given these calamitous drawbacks, before trimming, researchers should provide some evidence that data errors exist and that trimming is successful at eliminating such errors. Otherwise, trimming can either spuriously induce evidence against the efficient market hypothesis or against behavioral explanations.

REFERENCES

Ali, A., L. Hwang, and M. Trombley. 2000. Accruals and future stock returns: Tests of the naïve investor hypothesis. *Journal of Accounting, Auditing and Finance* 15 (2): 161–181.

Ali, A., X. Chen, and T. T. Yao. 2008. Do mutual funds profit from the accruals anomaly? *Journal of Accounting Research* 46 (2): 1–26.

Beneish, M., C. Lee, and R. Tarpley. 2001. Contextual fundamental analysis through the prediction of extreme returns. *Review of Accounting Studies* 6 (2–3): 165–189.

Bradshaw, M., S. Richardson, and R. Sloan. 2000. Do analysts and auditors use information in accruals? *Journal of Accounting Research* 39 (1): 45–74.

Campbell, J., L. Andrew, and C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Campbell, J., and T. Vuolteenaho. 2004. Bad beta, good beta. *American Economic Review* 94 (5): 1249–1275.

Collins, D., J. Gong. and P. Hribar. 2003. Investor sophistication and mispricing of accruals. *Review of Accounting Studies* 8 (2–3): 251–276.

Core, J. 2006. Discussion of an analysis of the theories and explanations offered for the mispricing of accruals and accrual components. *Journal of Accounting Research* 44 (2): 341–350.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52 (3): 1035–1058.

Daniel, K., and S. Titman. 1997. Evidence on the characteristics of cross-sectional variation in common stock returns. *Journal of Finance* 52 (1): 1–33.

DeBondt, W., and R. Thaler. 1985. Does the stock market overreact? *Journal of Finance* 40 (3): 793–805.

Dechow, P., and W. Ge. 2006. The persistence of earnings and cash flows and the role of special items: Implications for the accrual anomaly. *Review of Accounting Studies* 11 (2–3): 253–296.

Dechow, P., S. Richardson, and R. Sloan. 2008. The persistence and pricing of the cash component of earnings. *Journal of Accounting Research* 46 (3): 537–566.

Desai, H., S. Rajgopal, and M. Venkatachalam. 2004. Value-glamour and accruals mispricing: One anomaly or two? *The Accounting Review* 79 (2): 355–385.

Dopuch, N., Seethamraju, and W. Xu. 2010. The pricing of accruals for profit and loss firms. *Review of Quantitative Finance and Accounting* 34 (4): 505–516.

Fairfield, P., J. Whisenant, and T. Yohn. 2003. Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review* 78 (1): 353–371.

Fama, E., and J. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–636.

Fama, E., and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (2): 427–465.

Fama, E. 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49 (3): 283–306.

Gow, I., G. Ormazabal, and D. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.

Hackel, S., J. Livnat, and A. Rai. 2000. A free cash flow investment anomaly. *Journal of Accounting, Auditing and Finance* 15 (1): 1–24.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (2): 125–153.

Hirshleifer, D., K. Hou, S. Teoh, and Y. Zhang. 2004. Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38 (December): 297–331.

Hirshleifer, D., K. Hou, S. Teoh, and Y. Zhang. 2010. *The Accrual Anomaly: Risk or Mispricing. Management Science*, forthcoming.

Hribar, P., and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (1): 105–134.

Jegadeesh, N. 1990. Evidence of predictable behavior of security returns. *Journal of Finance* 45 (3): 881–898.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48 (1): 65–91.

Klein, A., and C. Marquardt. 2006. Fundamentals and accounting losses. *The Accounting Review* 81 (1): 179–206.

Knez, P. M., and M. J. Ready. 1997. On the robustness of size and book-to-market in cross-sectional regressions. *Journal of Finance* 101 (4): 1355–1382.

Kothari, S., J. Sabino, and T. Zach. 2005. Implications of survival and data trimming for tests of market efficiency. *Journal of Accounting and Economics* 39 (1): 129–161.

Kraft, A., A. Leone, and C. Wasley. 2006. An analysis of the theories and explanations offered for the mispricing of accruals and accrual components. *Journal of Accounting Research* 44 (2): 297–339.

Levi, S. 2008. Voluntary disclosure of accruals in earnings press releases and the pricing of accruals. *Review of Accounting Studies* 13 (1): 1–21.

Louis, H., D. Robinson, and A. Sbaraglia. 2008. An integrated analysis of the association between accrual disclosure and the abnormal accrual anomaly. *Review of Accounting Studies* 13 (1): 23–54.

Mishkin, F. 1983. *A Rational Expectations Approach to Macroeconometrics: Testing Policy Effectiveness and Efficient Markets Models*. Chicago, IL: University of Chicago Press for the National Bureau of Economic Research.

Narayanamoorthy, G. 2006. Conservatism and cross-sectional variation in the post–earnings announcement drift. *Journal of Accounting Research* 44 (4): 763–789.

Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.

Richardson, S., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437–485.

Shevlin, T., C. Mashruwala, and S. Rajgopal. 2006. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics* 42 (1–2): 3–33.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Teoh, S., I. Welch, and T. J. Wong. 1998a. Earnings management and the long run market performance of the initial public offering. *Journal of Finance* 53 (6): 1935–1974.

Teoh, S., I. Welch, and T. J. Wong. 1998b. Earnings management and the underperformance of seasoned equity offerings. *Journal of Financial Economics* 50 (1): 63–99.

Teoh, S., and T. J. Wong. 2002. Why new issue and high-accrual firms underperform: The role of analysts' credulity. *Review of Financial Studies* 15 (3): 869–900.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76 (3): 357–373.

Zach, T. 2003. *Inside the "Accrual Anomaly."* Working paper, University of Washington.

Zhang, H., and L. Shi. 2007. *Can the Earnings Fixation Hypothesis Explain the Accrual Anomaly?* Working paper, Nanyang Technological University.

THE ACCOUNTING REVIEW
Vol. 86, No. 4
2011
pp. 1477–1485

American Accounting Association
DOI: 10.2308/accr-10045

BOOK REVIEWS

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Business, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

KRISLERT SAMPHANTHARAK and ROBERT M. TOWNSEND, *Households as Corporate Firms: An Analysis of Household Finance Using Integrated Household Surveys and Corporate Financial Accounting*, Econometric Society Monographs (Cambridge, U.K.: Cambridge University Press, 2010, ISBN 978-0-521-19582-9, pp. xii, 199).

This Econometric Society monograph is a result of over ten years of research by the authors on household finance in developing countries. They lay out the basic approach and build on and use the promulgated financial accounting standards for the corporate world (private sector) to construct three primary financial statements for households. These statements—the balance sheet, income statement, and statement of cash flows—use panel data from integrated household surveys in developing countries. The objective of this study is to develop a conceptual framework that will help researchers define and efficiently measure household productivity, risk, liquidity management, and other financial situations that families encounter. These household surveys are essential, as they enhance researchers' understanding of the behavior of households in developing countries and for their governments to evaluate and formulate policies. This is a path-breaking study because, although there have been other household surveys in developing countries, the wording or meaning of the questions in previous surveys has been sometimes inconsistent or unclear, and the concepts have not been well defined, which has led to measurement errors. The principal contribution of this monograph is the conceptualization of the accounts and a clear definition of the variables of interest, which helped improve the clarity of the survey questionnaire, and a systematic treatment of the issues.

The authors present the challenges and solutions to effectively construct financial statements from integrated household surveys. They argue that there is a need to impose an accounting framework on the survey data which will allow researchers to link the study of household finance at the micro level to the macro economy. The major advantages of this accounting framework are that it helps the researcher define the variables, and that the three financial statements provide a simple way to apply standard financial accounting to the study of household finance.

The authors clearly explain that the use of high-frequency household surveys is necessary for the purpose of this project to be realized, and they choose to use data from the Townsend Thai Monthly Survey. This survey deliberately selects two distinctive households with both typical and nontraditional challenging transactions. These surveys provide the basis for comparative advantages and the treatment of unusual transactions. Because case studies are specific and cannot be easily generalized, the authors supplement each finding from these two households with the quartiles from their corresponding provinces.

The eight distinct key findings from the Townsend Thai Monthly Survey are as follows. First, the survey revealed a large dispersion of average rates of return on assets (ROA) across households. Second, the differences between ROA and return on wealth or equity (ROE) can differ substantially between households.

Third, ROA drops dramatically when the imputed opportunity cost of household labor is accounted for. Fourth, income and cash flow have high volatility, but consumption is relatively smoother. Fifth, some households appear to base financial and consumption behaviors on accrued net income rather than cash flow. Sixth, cash finances consumption and investment-cash flow deficits. Seventh, increases in household wealth are associated with cash increases in more developed areas and do not correspond to households in less developed areas. Eighth, investment-cash flow sensitivity analysis revealed that the rural and semi-urban households in the sample seem to face a liquidity constraint, but the reverse may not necessarily hold.

This monograph is divided into three parts. Part I, "Households as Corporate Firms," introduces the topic and the objective of the study, and provides a conceptual framework of the study. Part II, "Household Financial Accounting," presents the actual construction of household financial statements from household surveys. Part III, "Household Finance," presents the financial analysis of the two case studies, gives an application of the accounts to the economic modeling of household finance, discusses the lessons learned from this study, suggests ways to improve the design of future surveys, and presents some limitations. The Appendix presents the financial statements constructed for the two case studies.

Chapter 1 proposes the conceptual framework that the authors plan to use in order to analyze households as corporate firms. Corporate financial accounting is used to build upon, and the accounts are modified to appropriately fit the very diverse transactions for households in developing countries. This model is then imposed on data from integrated household surveys to construct the balance sheet, income statement, and statement of cash flows. The challenges that arise from modeling households in developing countries include the inseparability of consumption and investment, distinguishing between income and cash flow, and developing definitions of income, multi-period production, and consumption of household production.

Chapter 2 provides a conceptual framework. It presents the analogy of households as corporate firms, which serves as the framework for constructing household financial accounts. Unlike the corporate firm, which is a collection of assets, a household consists of a collection of individuals. Even though households contrast to corporate firms in some important ways, conceptually they are very similar. The focus in this monograph is on the accounts that are consolidated and aggregated over all activities in the household.

Part II of the monograph provides an in-depth explanation of the concept of creating corporate accounts from household finance and its limitations. The data needed to convert the corporate accounting structure to the household financial structure must possess certain characteristics, such as panel data collected at a high frequency, and it must cover a wide range of activities performed by households, while remaining detailed enough to identify individual transactions. Chapter 3 gives an overview of the household surveys (LSMS and Financial Diaries Project) used in other studies and their limitations for use in this research. It explains the authors' decision to utilize the integrated panel survey of households in the Townsend Thai Monthly Survey. The Townsend Thai Monthly Survey conducted in rural and semi-urban areas of Thailand is made up of four parts: an initial census, a baseline interview, forms that gather information on formal and informal institutions, and monthly interviews. The two major benefits of this survey are the level of detail and the frequency of the interview at which the data are collected, which is very valuable to the study of households as corporate firms. These high frequency data are "necessary for the analysis of liquidity, the short-term smoothing of consumption, the protection of investment from cash flow fluctuations, and the financing of cash flow budget deficits" (p. 38). The project used surveys to gather reliable information on the formal and informal mechanisms used in each household for financial support.

Chapter 4 explains the construction of household financial statements from a household survey. The primary purpose of the experiment is to conceptualize the accounts, which is the most complex and problematic, and necessitates some choices and adaptations when it comes to unusual transactions. In order to offer what little uniformity they can, the authors use explicit definitions of the items in each account and explain the assumptions made to define and measure the various household transactions that do not conform to the corporate model. The authors argue that the accounting framework allows them to remain accurate while being flexible. The problem does not stem from the framework itself, but instead from the comparability from household to household in order to analyze behavior.

Part III discusses how the accounts are used in an analysis of household finance. Chapter 5 focuses on a case study of two households chosen from the Townsend Thai Monthly Survey. The reason the case study approach is selected by the authors is that it is often used in corporate financial analysis by creditors and analysts alike in order to find out the state of a firm and whether that performance can be said to be strong or poor on an objective scale. The two distinct types of households selected for this study are prototypical examples. The first household, household A, is in a relatively well developed region, and the household itself is very rich. The main occupations of the households are trading animal feed and raising milk cows. The husband works as a government employee and helps the wife. The second household, household B, is from a relatively less developed region, and the household is relatively poor. There are five people in this household, and all household members (except for the baby) work primarily as unpaid workers in the household's rice cultivation. The three financial statements constructed for these two households are presented in the Appendix.

Household productivity is determined by using standard corporate financial ratios and rates of return: first the ROA, and second the ROE. The authors choose to gather a risk-adjusted return figure—a return not driven by production technology riskiness—for household production comparisons. They turn to the Capital Asset Pricing Model (CAPM) for guidelines on measuring the price of risk for households. Using traditional financial gauges such as alpha and beta for risk and risk management, it is determined that household A has little correlation to the market—the village—with a flat beta, and household B's beta exceeds 1, meaning that its return on tangible assets moves with the village but in greater magnitude. The authors also examine the fluctuation of net income, cash flow, consumption, and investment by analyzing the co-movement of these variables. Because there are only two households surveyed in this particular application, the data gathered and presented are supplemented by quartiles from the provinces associated with the households.

Chapter 6 presents an application of the constructed financial accounts to an analysis of other rural households in the survey. It uses these accounts to analyze household liquidity constraints and the financing of household investment in fixed assets. The authors discuss the relevant corporate finance literature on investment and its financing and draw the parallels from corporate firms to households. Most of the data used in the empirical analysis in this chapter are from the Townsend Thai Monthly Survey and include 531 households surveyed over an 84-month period (7 years). The authors again stress the usefulness of the detailed data items for the analysis of investment financing and emphasize that high-frequency data are very relevant for the study of liquidity constraints. The demographic information gathered from the questionnaire and the kinship variables serves an important purpose in this analysis, because it can help reduce liquidity constraints for households in developing countries. Descriptive statistics on the initial fixed asset size for these households, positive household investments, median household investment value, and median investment rate are presented and discussed to highlight the impact of financial investments and financing fixed investments through various financial devices on households at various economic levels.

Chapter 7 concludes this monograph by discussing the various issues of measurement and modeling of households as corporate firms. It presents the lessons learned by discussing the issues faced and the decisions made to resolve these challenges. It also presents the limitations of the study and suggests ways to improve future surveys for other researchers.

I thoroughly enjoyed studying this monograph, which is truly a path-breaking research that was badly needed in the field of household finance in developing economies. This monograph will be very influential and useful in the field of microfinance, where the microfinance institutions are always evaluating household finance for every potential client before a microloan is made. The strength of this research study stems from:

(1) the explanation of the conceptual framework: it gives solid insight into what the authors hoped to accomplish and explains in detail the logistics of the study;
(2) the very clear definition of each variable in the financial statements so that they are consistent with each other and consistent with the meaning implied in economic theory;
(3) the conceptualization of the accounts which necessitated making some hard decisions for unusual transactions;
(4) the measurement of these variables is very precise, lending credibility to the results;
(5) the level of detail in the survey questionnaire, which allowed the interviewer to ask about each different formal and informal mechanism individually for a thorough study of the role which these devices have in liquidity, financing, and investment and portfolio management (p. 34);
(6) the frequency with which the data were collected is the key for short-term analysis of household behavior which cannot be accomplished by conducting even quarterly interviews; most households in developing economies financially survive on a week-to-week basis;
(7) its collection of demographic and environment variables and the importance of kinship which can finance and lessen the liquidity constraints of a household;

(8) the experiment is well thought out, the data are systematically organized, the analysis is rigorous, and the results are well presented and discussed;
(9) suggestions are made for improving this survey in the future by other researchers; and last and most critical;
(10) it has demonstrated to the world that household finance of families in the lower socio-economic strata can be studied.

However, like all research, this study also has its limitations. The authors have listed several constraints of households in applying the formal accounting principles. Data collection was a daunting task, and this study also has limitations relating to the level of detail in the questionnaire which may increase the chance of measurement errors. It is also more demanding of time and imposes higher monetary costs, as well as higher opportunity costs for the households involved. A few problems unique to the Townsend survey also have been complex timing issues that can skew the data. Another major issue in the application of this experiment is in its ability to be cross-applied to other countries on a usable scale. One problem with the monograph is that it is written in professional jargon, and hence the average person would have difficulty understanding it. A good example, in addition to the language barrier placed on the common reader, is the issue with applying the accounts themselves on a global scale, when many transactions in the experiment are unique to households in developing countries.

In summary, I recommend that this monograph be required reading for all students interested in the field of social entrepreneurship, loan officers in microfinance institutions, not-for-profit and nongovernmental organizations, and researchers in this field. I plan to apply the concepts learned as I embark shortly on implementing a microfinance project in sub-Saharan Africa.

SALEHA B. KHUMAWALA
Associate Professor of Accounting
University of Houston

HENRI FORTIN, ANA CRISTINA HIRATA BARROS, and KIT CUTLER, *Accounting for Growth in Latin America and the Caribbean—Improving Corporate Financial Reporting to Support Regional Economic Development* (Washington DC: The World Bank, 2010, ISBN 978-0-8213-8108-3, available online, pp. xxxi, 164).

This book, written by three authors from The World Bank and edited by the same institution, describes the status and development of accounting and the accounting and auditing professions in Latin America and the Caribbean. It categorizes the very wide range of accounting and auditing development among these countries. The authors report on The World Bank's efforts to develop this region and the capability of these professionals, as well as to develop the region's private economic sector. They also attempt to show how the development so far in this region differs from others on the planet.

The book was published in 2010, but it was completed during the global financial crisis of 2007/09. Obviously, it does not report some very big changes in this region which occurred in several countries in 2009 and 2010. It is heavily based on The World Bank's ROSC (Reports on the Observance of Standards and Codes) project, as part of an almost ten-year-old program in which the bank investigates the region's *status quo* regarding IFRS acceptance, evolution of the profession, and the evolution of transparency, accountability, and governance.

The authors very interestingly and repeatedly assert that the imposition of accounting and auditing rules without corresponding enforcement actions ends up as merely "virtual rules," which is quite typical in some of these and other developing countries. That is why the book frequently calls attention to the gap between the standards issued and the effectiveness of actual compliance with them. It also calls attention to situations, common in these and other countries, in which multiple standard-setting and sometimes supervisory entities exist, fragmenting efforts and wasting financial and human resources.

The work is not purely descriptive, though, as it defends countless concepts related to the usefulness of good accounting and independent auditing for economic development, the benefits of regional integration, the creation of a better business climate when the market receives high-quality information, etc. On the one hand, the book attempts to depict the regional situation, yet, on the other, it seeks to promote desirable change.

Another finding which permeates this work is that the fact that the regional economy mainly comprises strongly family-dominated companies, followed by state-owned companies and, ranking third only, by public companies, strongly gives rise to resistance against information disclosure (mainly by family-owned companies—afraid of informing the competition) and the dissemination of good information by the government itself (although exceptions exist).

As already stated, some extraordinary changes in the region occurred after the book was completed. To give an example, in 2010, all Brazilian companies are adopting IFRSs or IFRSs for SMEs (small and medium enterprises), in consolidated and individual (i.e., parent and subsidiary company) financial statements, except for banks and insurance brokers, which limit this adoption to the consolidated financial statements only. Various tables in the book on the countries' specific situations regarding different requirements are also outdated, although this is due to the speed with which changes have been taking place in the region.

As for banks and insurance brokers, the book shows that, as a result of public regulators' fairly good supervisory structure, the quality of their financial statements usually exceeds, on balance, the quality in other companies. And the authors quite strongly criticize the lack of specific and simplified accounting standards for SMEs, which is correct, as studies on the adoption of the IFRSs for SMEs are only starting now (with Brazil taking the lead, as already indicated).

When the book mentions state-owned enterprises, the authors make numerous comments on the deficiencies in their financial statements. This assertion is somewhat incomplete, however, as these companies tend to be leaders in terms of financial statement quality in some countries, including Brazil, although they really are anything but transparent in other countries.

Chapter 1 provides an introduction and the background needed to understand the problem under discussion, mentioning several of the abovementioned assertions. Chapter 2 describes the national corporate financial reporting frameworks, highlighting the cultural problem of Roman code law in most countries in the region, as well as the legalistic and formalist tradition (as opposed to common law, which exists only in some Caribbean countries), the role of individual financial statements in commercial law (consolidated financial statements are considered mere complementary information and are not used on a large scale), the less than widespread use of MD&A (Management's Discussion & Analysis), etc. The link, broken in Brazil only, between financial and tax accounting has led to the practice of trying to diminish the company's tax burden, and through other than sound accounting practices. In the region, individual financial statements play a very large role as, according to the book itself, they underlie not only taxation (except in Brazil) but also dictate the required dividend distributions as well as determine stock values when buying back shares from dissident stockholders, etc.

The information that is missing is that, together with Italy, Brazil is taking the lead in the adoption of IFRSs for individual financial statements. Also in Brazil, a complete distinction is being drawn between accounting for information disclosure purposes and accounting for tax purposes, and, for the first time, the accounting standards are being adopted without interference from the income tax authority. In the other countries, however, tax influence remains very strong, so that managements prepare their individual financial statements with a focus on tax consequences, as the book correctly affirms. So, in these countries it is possible to apply IFRSs only in the consolidated financial statements, which oblige the enterprises to maintain two sets of books.

Chapter 3 discusses the role and work done in the good name of the accounting profession. The book asserts that, generally speaking, the accounting profession has resisted efforts to assume additional obligations, mainly to avoid the necessity of having to set accounting standards. As for the accounting profession, in some countries small local firms dominate the auditing sector, and often act in overt opposition to the big international firms. Auditors still maintain their self-regulatory power, but some countries are moving toward the adoption of IFAC's international auditing standards. The authors describe the current state of affairs in the conduct of the auditing profession: the status of peer review, the mandatory rotation of auditors, and the still very incipient creation of regulatory entities like the PCAOB.

Chapter 4 discusses the role of education for accounting, evidencing the gap between the profession's needs and what is currently offered by the universities. Chapter 5, on the adoption of IFRSs, is the most outdated, as mentioned, but it highlights the role of the Comitê de Pronunciamentos Contábeis (CPC)—the Accounting Pronouncements Committee, which was created in Brazil by private entities (representing auditors, accountants, public companies, academics, analysts, the stock exchange), independently of the State, but in which governmental entities participate as observers, with the right to speak but not to vote (the Brazilian SEC, the Central Bank, the Internal Revenue Service, and the Insurance Companies Supervisory Board), in addition to other entities that represent banks and industries.

In Chapter 6, the authors voice concerns over compliance with both accounting and auditing standards; this concern exists because the banking system does not use, as strongly as it should, financial statements as the basis for its credit operations. This creates a huge gap in the financial statement quality between companies traded in the stock market and others which are not. Chapter 7 proposes a roadmap for reform, quite firmly emphasizing the need to adopt special rules for SMEs, the modernization of the academic curriculum, and the creation of an audit oversight board independently of the auditing profession.

Appendices describe the role of The World Bank's ROSC program, its regional coverage, the adoption of IFRSs, and the development of a strategic view for the necessary reforms.

Despite some inevitably outdated aspects, the authors provide an excellent overview of what is going on with accounting and the accounting profession in Latin America and the Caribbean. Thus, it is useful not only for academic purposes, but also for investors and companies with economic interests in the region.

ELISEU MARTINS
Professor of Accounting
Universidade de São Paulo

STEPHEN A. ZEFF, *Insights from Accounting History: Selected Writings of Stephen Zeff* (New York, NY: Routledge, 2010, ISBN: 978-0-415-55429-9, pp. xxxiii, 522).[1]

This is an appropriate juncture to be reviewing a work on insights from accounting history, given the pressing need for accounting academics in the U.S. and internationally to rediscover the potential of history for gaining a fuller understanding of accounting practice and the environment in which it operates and which it has helped to shape. This was the view taken by the American Accounting Association (1970), whose report acted as a clarion call for American educators, with the support of the standard setter, to raise the profile of accounting history. Forty years on, the promise of an educational and research agenda informed by insights from accounting history has by and large faltered in U.S. schools, where accounting research has become dominated by ahistorical research methodologies (Mouck 1995; Lee 2006; Vangermeersch 2008). The situation persists despite the commitment of American accounting historians to rectifying the situation, but who find themselves obliged to seek publication outlets outside the mainstream U.S. journals. This compilation of historical writings by Stephen Zeff is unusual in that as many as six of the 15 post-1980 articles included in the collection were originally published in generalist American accounting journals, including two in *The Accounting Review*—proportions which, although seemingly low, bear witness to Zeff's high standing in the U.S. accounting academy.

The retreat from history by the mainstream is a wasted opportunity that this book, and the Routledge series on *Historical Perspectives in Accounting* of which it forms part, is attempting to address. Among its benefits, accounting history can show us the roots of present practice and help us to better understand human nature, better understand present social, economic, and organizational structures, be aware that the problems of the present are not unique, and finally give us a fuller appreciation of the patterns of regulation and abuse, including the likely reactions of the major players (Jones and Oldroyd 2009). All of these themes are explored in the book, which, in keeping with the rest of the Routledge series, focuses on the contemporary evolution of accounting thought, practice, and standard setting. Zeff's work is therefore highly pertinent to the challenges facing the modern world of accounting.

---

[1] The arrangements for this review were made by Professor Richard Macve, of the London School of Economics.

The book comprises a foreword by the distinguished accounting historian, J. R. Edwards, an introduction by the author, and 18 papers that were originally published between 1962 and 2007. The foreword places the work in its historiographical context, examines Zeff's research methods, and comments on his academic career. As Edwards observes, the collection does not fully reflect Zeff's historical research, as it necessarily omits his books and monographs. These are listed in an appendix. The introduction provides an overview of the collection, discusses the genesis of the various papers, and reveals their motivation.

Zeff identifies three broad themes running through the papers: "the historical development of the process of setting accounting standards, the evolution of accounting thought, and the evolution of the accountancy profession and of audit practice" (p. xxxii). Zeff's career-long interest in the standard setting process, especially the political lobbying involved, is signaled by the second paper in the collection, "The Rise of 'Economic Consequences'," published in 1978 as part of the *Stanford Lectures in Accounting*. Here Zeff is referring to a growing realization by interest groups that accounting standards could produce economic consequences that they in turn could attempt to influence through lobbying (p. 17). The paper relates how such interventions affected the work of the Committee on Accounting Procedure (CAP), the Accounting Principles Board (APB), and the Financial Accounting Standards Board (FASB) in a number of instances between 1941and 1977. According to Zeff, an abridged version of the paper published in the *Journal of Accountancy* is the most widely cited article of his career (p. xxii).

The evolution of accounting thought is the theme of the first paper in the book which discusses replacement cost theory in the context of contemporary business practice and earlier 20th century writers on price-level accounting. Other theoretical issues discussed in the collection include Kenneth MacNeal's advocacy of market-based valuations and the evolution of the conceptual framework in the U.S., which Zeff traces back to Paton and Canning in the 1920s. The impact of Canning's *magnum opus* on accounting theory, *The Economics of Accountancy*, is assessed in a separate paper.

The evolution of the accountancy profession and audit practice is the final theme identified in the introduction. The establishment of accounting principles in the U.S. from 1917 to 1972 is linked to the profession's fear of government interference (p. 126). The decline in advocacy writing on standard setting and accounting principles by the Big 8 firms in the 1970s is explained in terms of a number of factors, including the loss of flexibility in the interpretation of GAAP that occurred as regulation became more systematic (p. 143). The earlier willingness of the profession to take a lead in matters of principle is explored further in a study explaining Arthur Andersen & Co.'s decision in 1946 to decouple "present fairly" and "conformity with generally accepted accounting principles" in the wording of their audit report (pp. 157–160). Indeed, looking back over his life's work, one of Zeff's main conclusions is the weakening in "the vitality and dedication to the public interest by the U.S. accountancy profession" occurring since the 1970s (p. xxxiii). This decline in professional values is explored in more detail in a two-part article analyzing the roots of the U.S. accountancy profession's transition from "professional firms that happened to be businesses" pre-1970 to "businesses that happened to render professional services" in the Enron era. Zeff recognizes a number of changes at work in the market for accountancy services and the regulatory environment, and identifies "the root cause of the questionable decisions attributable to audit firms in recent years" as the incentives given to their partners (p. 437). This is a good illustration of the value of history, as without a comprehension of the situation as it once was, how could one appreciate that a change had taken place and when, let alone understand the nature of the change and the causes? Studies relating to particular organizations include an explanation of DuPont's policy from 1910 to 1950 of rotating audit firms and an appraisal of the reporting changes of Unilever in the 1940s and their impact on U.K. regulation after the Second World War. In short, this is an interesting and eclectic collection of papers that it is difficult to do full justice to in a review.

Finally, if one of the lessons to be learned from the book is that history can provide useful insights into the world of accounting, another is of the value of engagement with professionals. A division now exists between professional and academic accountants whereby much of mainstream accounting research has come to be regarded as incomprehensible and irrelevant by the former, in contrast to the "vibrant dialogue" that once existed between the two communities (pp. 400, 402). The level of disengagement is compounded at the "critical" end of the academic research spectrum, where there is a tendency to disparage the profession as representing the antithesis of social justice. By contrast, Zeff's *Insights from Accounting History* is premised on the value of engagement, that the best way to uncover the facts is to persuade the people in the know either to speak to you or to make their papers available. Edwards comments on this aspect of the work in the foreword (p. viii), an approach that has gained Zeff access to privileged information relating to the AAA, CAP, APB, and Securities and Exchange Commission, among others (pp. xviii, xxi, xxv). There is a very valuable lesson for accounting researchers to learn here.

REFERENCES

American Accounting Association (AAA). 1970. Report of the committee on accounting history. *The Accounting Review* 45 (Supplement): 52–64.

Jones, M. J., and D. Oldroyd. 2009. Financial accounting: Past, present and future. *Accounting Forum* 33 (1): 1–10.

Lee, T. A. 2006. The war of the sidewardly mobile corporate financial report. *Critical Perspectives on Accounting* 17: 419–455.

Mouck, T. 1995. Irving Fisher and the mechanistic character of twentieth century accounting thought. *Accounting Historians Journal* 22 (2): 44–83.

Vangermeersch, R. 2008. The Academy of Accounting Historians 1973-2008-2023. *Accounting Historians Notebook* 31 (2): 1–8.

DAVID OLDROYD
Professor of Accounting
Newcastle University Business School

CAPSULE COMMENTARIES

By the Book Review Editor

LOUISE CRAWFORD, CHRISTINE HELLIAR, and DAVID POWER, *Politics or Accounting Principles? Why Was IFRS 8 So Controversial?* Briefing (London, U.K.: ICAEW Centre for Business Performance, 2010, ISBN 978-0-85760-056-1, pp. 31).

IFRS 8, on segment reporting, has been, so far, the only IASB standard whose endorsement for use in the European Union has been taken up by the European Parliament. It requires the reporting of operating segments that are used internally by the company when it makes operating decisions. IFRS 8, which was issued in November 2006, mirrored the Financial Accounting Standards Board's SFAS 131 and thus was viewed as an importation of U.S. GAAP into Europe via the IASB, a politically charged argument. Yet news articles in *Accountancy Age* reported that its substantive content, wholly apart from its provenance, was controversial. Was it politics or a dispute over proper accounting that explained the interest by a committee of the Parliament?

The researchers' aim was to learn whether a sample of stakeholders believes that this special attention by the Parliament was because the standard's substantive content was controversial or because, instead, the vocal opposition was politically motivated. This research team from the University of Dundee conducted semi-structured interviews with 15 key stakeholders, including preparers, regulators, legislators, auditors, institutional investors, and users. Most of the booklet is devoted to an elaborate and interesting discussion of their findings.

Perhaps to the authors' surprise, most of the stakeholders did not believe that the substance of the standard was all that controversial. The authors therefore concluded that one or more of several politically inspired motivations seemed to be the explanation. One hopes that this briefing will be followed by other such succinct studies of the reception accorded the IASB's more noteworthy standards.

FRANK CLARKE, GRAEME DEAN, and MURRAY WELLS, *The Sydney School of Accounting: The Chambers Years* (Sydney, N.S.W., Australia: The University of Sydney, 2010, ISBN 978-1-74210-203-0, pp. 136).

This reminiscence was written to celebrate the 50 years since Ray Chambers (1917–1999) became the Foundation Professor of Accounting at The University of Sydney in 1960. For 40 of those years, Chambers was the intellectual leader and an always stimulating presence: a rigorous theorist and prolific author whose impact was felt around the world, the founder and first editor of a major accounting research journal, an indefatigable correspondent, the builder of an innovative accounting curriculum which established accounting as a proper university discipline, and an academic who genuinely sought a dialogue with the accounting profession.

The chapter on Chambers' efforts to gain a hearing for his proposal for reform known as Continuously Contemporary Accounting, or CoCoA, usefully fills some gaps in the literature on standard setting in Australia during the inflationary 1970s. It is also useful to have a reprise of Murray Wells' review of the obstacles that Chambers overcame in his efforts to launch *Abacus* in 1965. The tale of how much of Chambers' incredibly rich trove of books and papers was saved at the last minute from mindless destruction and then was resourcefully converted into an Archive for use by researchers is well told.

The sense of community and vitality within the accounting group during the five decades of what has come to be known as the Sydney School of Accounting is well conveyed by the three authors, who drew with good effect on their recollections as insiders.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature. Topical areas of interest to the journal include accounting information systems, auditing and assurance services, financial accounting, management accounting, taxation, and all other areas of accounting, broadly defined. The journal is also open to all rigorous research methods.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and editors are responsible for providing constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary.*

FORMAT

1. All manuscripts should be formatted in 12-point font on 8 1/2 × 11″ paper and should be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins should be at least one inch from top, bottom, and sides.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we".
5. A cover page should show the title of the paper, all authors' names, titles and affiliations, email addresses, and any acknowledgments.
6. The American Accounting Association encourages use of gender-neutral language in its publications.

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years.* All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a cross-sectional equation; re-form.* See *Webster's* for correct usage.

Keywords: The abstract must be followed by at least three keywords to assist in indexing the paper and identifying qualified reviewers.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words (150 maximum) should be presented on a separate page immediately preceeding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. The Keywords statement should appear immediately below the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. When the reference list contains two or more works by the same author (as the only author or first of two or more authors) published in the same year, the suffix a, b, etc., is appended to the date in the within-text citations and in the "References" section. For example, (Johansson 2004a, 2004b, 2004c; Baiman and Rajan 2002a, 2002b; Dhaliwal et al. 2005a, 2005b).
4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works should use acronyms or short titles where practicable. For example: (NCFFR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in *alphabetical order* according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Use authors' initials instead of proper names.
3. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
4. Date of publication follows the author's (authors') name(s).
5. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

Baiman, S., and M. Rajan. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.

Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: http://hsgac.senate.gov/files/111803berry.pdf.

Cohen, D., A. Dey, and T. Lys. 2005. The Sarbanes Oxley Act of 2002: *Implications for Compensation Structure and Risk-Taking Incentives of CEOs*. Working paper, New York University, University of Chicago, and Northwestern University.

Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition*. Ann Arbor, MI: University of Michigan.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

Dechow, P. M., S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

Dechow, P. M., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.

Dhaliwal, D., Erickson, and O. Li. 2005a. Shareholder income taxes and the relation between earnings and returns. *Contemporary Accounting Research* 22: 587–616.

Dhaliwal, D., L. Krull, O. Li, and W. Moser. 2005b. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43: 675–708.

Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.

Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).

Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109*. FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Boston, MA: Houghton Mifflin.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.

Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics* (Winter) 30: 555–574.

National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting*. Washington, D.C.: NCFFR.

Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.

Nelson, M. W. 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy*, edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.

Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.

Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports*. Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be inserted using the "footnote" or "endnote" feature of the word processing software which will automatically number the footnotes throughout the manuscript with superscript Arabic numerals.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state upon submission that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. New manuscripts must be submitted through the Manuscript Submission and Peer Review System, located at http://accr.allentrack.net. The site contains detailed instructions regarding the preparation of files for submission. Revisions of manuscripts originally submitted before June 1, 2011 should be sent via email to the incoming senior editor, John Harry Evans III, at tar@katz.pitt.edu. Please submit separate files for (1) the manuscript's title page with identifying information (not forwarded to reviewers), (2) the manuscript with title page and all other identifying information removed, and (3) any necessary supplemental files, such as experimental instructions and/or response memoranda on invited revisions.
4. A nonrefundable submission fee is required unless explicitly waived by the editor for invited revisions of previous submissions. The submission fee in U.S. funds is $200.00 for members and $400.00 for nonmembers of the AAA payable by credit card (VISA or MasterCard only). The payment form is available online at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If you are unable to pay by credit card or have any questions, please contact the AAA Member Services Team at (941) 921-7747 or info@aaahq.org.
5. Revisions must be submitted within 12 months from the decision letter inviting a revision.

COMMENTS

The Accounting Review welcomes submissions of comments on previous *TAR* articles. Comments on articles previously published in *The Accounting Review* will generally be reviewed by two reviewers, usually including an author of the original article (to assist the editor in evaluating whether the submitted comment represents the prior article accurately) and an additional reviewer who is independent of the original article. If a comment is accepted for publication, the original author will generally be invited to reply. All other editorial requirements, as enumerated above, apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The AAA's Executive Committee policy (originally adopted in 1989, and amended in 2009) is that the objective of the Association-wide journals (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. To fulfill this objective, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles.

JULY 2011 PLACEMENT ADS

The deadline for free position ads to be included in this section is two months prior to the date of publication. *The Accounting Review* is published in January, March, May, July, September, and November. Position ads in the "Placement Ads" are free with the purchase of a job posting in the AAA Career Center. For more information on how to purchase a job posting in the Career Center and receive your free position ad in *The Accounting Review*, go to the AAA website at http://aaahq.org and click on "Career Center," or call our office at 941-921-7747.

SOUTHERN ILLINOIS UNIVERSITY CARBONDALE, School of Accountancy is accepting applications for a tenure-track position at the rank of Associate or Full Professor of Accounting starting August 16, 2011. Applicant must have an earned Doctorate in Accounting and a minimum of five years post Ph.D. teaching experience at the effective date of employment. Applicant must have a strong commitment to a quality research program and excellent teaching. An established research record in premier and high quality publications is expected. Research responsibilities will include working with doctoral students and chairing dissertations. Teaching responsibilities will include participation in the undergraduate, M.Acc., and Ph.D. programs. All areas of specialization will be considered. Submit letter of application, resume and contact information for three references by email to sobery@cba.siu.edu or by post mail to: Dr. Julie Sobery, Search Committee Chair, School of Accountancy, Mail Code 4631, Southern Illinois University Carbondale, 1025 Lincoln Drive, Carbondale, IL 62901. Telephone: (618) 453-2289, Fax (618) 453-1411. Review of applications will begin April 15, 2011 and continue until the position is filled. SIUC is an Affirmative Action/Equal Opportunity Employer that strives to enhance its ability to develop a diverse faculty and staff and to increase its potential to serve a diverse student population. All applications are welcomed and encouraged and will receive consideration.

CALIFORNIA STATE UNIVERSITY, LOS ANGELES invites applications for a full-time Assistant Professor position in the areas of International Accounting Standards, Audit/Fraud, Accounting Information Systems, Governmental Accounting, Tax Accounting, Financial, or Managerial Accounting. Start date is September, 2011. A doctoral degree must be from an accredited institution of higher education in accounting. Applicants must be academically qualified according to the AACSB standards. ABDs will be considered. Ph.D. or equivalent is required for retention after two years. Preference will be given to candidates with Ph.D. or equivalent at the time of employment. Please send a letter of application, *curriculum vitae*, three letters of recommendation, and official transcript from institution awarding highest degree to: Ms. Sue Bergstrom, College of Business and Economics, California State University, Los Angeles, 5151 State University Drive, Los Angeles, CA 90032-8120. An Equal Opportunity/Title IX Employer. Please see detailed description of position at: http://www.calstatela.edu/academic/position/csla_be.php

THE UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN invites nominations and applications for the position of Professor and Head of the Department of Accountancy in the College of Business. The candidate must possess an earned doctorate in Accountancy or equivalent and should give evidence of demonstrated excellence in research, teaching and service sufficient to merit the rank of Professor in Accountancy. Salary is competitive. The preferred starting date is on or before January 1, 2012, but is negotiable to accommodate the selected candidate. To receive full consideration, applications must be submitted online by September 1, 2011. Application materials must include a cover letter, *curriculum vitae*, and three letters of reference. Please create your candidate profile at https://jobs.illinois.edu Letters of recommendation may be submitted via mail or email to: Jean Seibold, Department of Accountancy, University of Illinois, 360 Wohlers Hall, 1206 S. Sixth Street, Champaign, IL 61820 or e-mail: accy@illinois.edu For further information please refer to http://www.business.illinois.edu/accountancy/recruithead.html or email accy@illinois.edu. The University of Illinois is an Affirmative Action/Equal Opportunity Employer and welcomes individuals with diverse backgrounds, experiences, and ideas who embrace and value diversity and inclusivity. (www.inclusiveillinois.illinois.edu)

MISSISSIPPI STATE UNIVERSITY, Adkerson School of Accountancy seeks applicants for a tenure-track position (Assistant/Associate/Full Professor) to begin Fall 2011. Rank is open; candidates must have a doctorate in accounting. Consideration will be given to ABD candidates. Candidates should have a strong commitment to publishing scholarly research. All areas of accounting will be considered. Professional certification and accounting practical experience are highly desirable. Salary and benefits are competitive. Position will remain open until filled. Send letter of interest, current *vitae*, and names of references (including contact information) to: James H. Scheiner, Director, Richard C. Adkerson School of Accountancy. Also, complete the Personal Data Information Form via the MSU On-Line Employment System at www.jobs.msstate.edu MSU is an AA/EOE.

THE UNIVERSITY OF PENNSYLVANIA, Wharton School's Accounting Department is seeking applicants for a full-time, tenure-track faculty position at any level: Assistant, Associate, or Full Professor, in Financial, Managerial, and Tax Accounting. The appointment is expected to begin July 1, 2012. Applicants must have a Ph.D. (expected completion by June 30, 2013 is acceptable) from an accredited institution and have an outstanding research record or potential. Demonstrated ability or potential for high quality research and teaching are important. Applications, including cover letter, *curriculum vitae*, and recent publications or working papers, must be submitted via our secure website: http://accounting.wharton.upenn.edu/. Junior-level candidates should also include contact information for up to three references who will be asked to submit letters of recommendation. The deadline for applications is November 30, 2011. Contact: Professor Brian Bushee, Accounting Department, The Wharton School University of Pennsylvania, 1300 Steinberg-Dietrich Hall Philadelphia, PA 19104-6365. The University of Pennsylvania values diversity and seeks talented students, faculty and staff from diverse backgrounds. The University of Pennsylvania is an equal opportunity, affirmative action employer. Women, minority candidates, veterans and individuals with disabilities are strongly encouraged to apply.

PEKING UNIVERSITY HSBC, Business School invites applications for a quality research and teaching position. A successful applicant should hold a Ph.D. or is expected to receive one before July 2012. Senior applicants should have excellent records on research and teaching. Located in the thriving entrepreneurial city of Shenzhen and adjacent to the financial hub of HK, Peking University is among the most prestigious education institutes in China. To apply, please submit your application package to hr_faculty@phbs.pku.edu.cn Your package should include a cover letter, a *curriculum vita*, a research paper, three reference letters, and evidences of effective teaching. Apply by July 31, 2011 to be considered for an interview at the AAA meeting in Denver, Co, August 6–10.

FAIRFIELD UNIVERSITY, Charles F. Dolan School of Business Accounting Department invites applications for a tenure-track position at the Assistant or Associate Professor level to begin January 2012 or Fall 2012. Candidates must have or be near completion of a doctoral degree. Professional certification in accounting is required. Qualified individuals in all specialty areas of accounting will be considered; preference will be given to those with an ability to teach in multiple areas. Candidates will be expected to demonstrate a strong commitment to excellence in teaching, scholarly research, and service to the department and the university. Fairfield University is Connecticut's Jesuit comprehensive university located on a scenic suburban campus near Long Island Sound about one hour outside New York City. The Charles F. Dolan School of Business is AACSB-accredited and was recently named a "top 15" accounting program by the Princeton Review's "Student Opinion Honors for Business Schools." Please send a letter of application, *curriculum vitae*, transcripts and three references to: Dr. Joan Van Hise, Chair; Accounting Department; Charles F. Dolan School of Business; Fairfield University; Fairfield, CT 06824; email: jvanhise@fairfield.edu. A review of candidates will begin immediately and continue until the position is filled or September 30, 2011, whichever occurs first. Fairfield University is an Affirmative Action/Equal Opportunity Employer. Women/minorities/persons with disabilities are strongly encouraged to apply.

THE UNIVERSITY OF DETROIT, Mercy's College of Business Administration seeks an Assistant Professor of Accounting. Ph.D. in accounting and active research agenda required; auditing/forensic accounting specialization desired. The College is accredited by AACSB and places major emphasis on teaching excellence. Michigan's largest, most comprehensive private university, the University of Detroit Mercy is an independent Catholic institution of higher education sponsored by the Religious Sisters of Mercy and the Society of Jesus. The university seeks qualified candidates who will contribute to the university's mission, diversity, and academic excellence. To interview at the American Accounting Association's conference in Denver, August 6-10, contact Dr. Robert Hutchinson, University of Detroit Mercy, 4001 W. McNichols Road, Detroit MI 48221, hutchirk@udmercy.edu, or (313)993-1161.The University of Detroit Mercy is an Equal Opportunity/Affirmative Action Employer with a diverse faculty and student body, and welcomes persons of all backgrounds. More information is available at http://udmercy.edu and http://business.udmercy.edu

WELLPOINT, Cincinnati, OH, seeks candidates for an Accounting and Analysis manager. The position requires a B.A./B.S. in accounting, finance, or business administration; 8-10 years of progressively more responsible experience in a senior financial management position with a medium to large insurance company; CPA, CFM, CMA or other relevant designation required; or any combination of education and experience, which would provide an equivalent background. M.B.A. or master's in accounting preferred. The manager, Accounting & Analysis, manages accounting and finance functions. Provides financial leadership, decision support and strategic direction. Position requires a process-oriented mindset with experience in streamlining processes, procedures and policies. Experience with PeopleSoft and Ariba systems. Strong supervisory skills are a must, with experience in developing and implementing departmental performance metrics. Experience managing offshore resources a plus. WellPoint is the nation's leading health benefits insurer and a Fortune Top 50 company. WellPoint is ranked as one of America's Most Admired Companies among health insurers by Fortune magazine and is a 2010 DiversityInc magazine Top 50 Company for Diversity. If interested, please apply directly via our website: http://www.careersatwellpoint.com/JobDescription.aspx?Identifier=48891, or go to www.wellpoint.com/careers and search for job number 48891 or job title. EOE

THE UNIVERSITY OF AKRON seeks candidates for a faculty position as Assistant/Associate/Professor of Accounting (Financial Reporting), teaching sophomore, junior, senior, and graduate courses, and engaging in research and scholarship activities. A Ph.D. in accounting from an AACSB-accredited institution or equivalent required. All applicants must be academically qualified under AACSB standards and must demonstrate a commitment to teaching excellence and to publishing high-quality research in peer-reviewed journals. Applicants must also demonstrate ability and commitment to contribute to the School's mission through service to the School, College, University, the profession, and the discipline. Preference shall be given to individuals with teaching and research interests in financial reporting and a related area. Professional certification is highly desirable. Also desirable are professional experience and ability to work closely with the professional accounting/tax community. For complete details and to apply for this position, Requisition 006563, visit: http://www.uakron.edu/about_ua/employment/job_openings. EEO/AA.

THE UNIVERSITY OF MINNESOTA DULUTH, The Department of Accounting in the Labovitz School of Business and Economics has a full-time, 9-month, tenure-track, Associate/Assistant/Instructor position available beginning Fall 2012. Position #171635. Essential qualifications include ABD (with a clear plan for completion by September 1, 2013) from a doctoral program in Accounting or Business Administration with a concentration in Accounting, from an AACSB-International accredited school or major internationally recognized university. Job duties and responsibilities include teaching, research and service. Individual will develop and teach accounting courses in auditing, as well as either financial or managerial accounting areas. For a complete position description and information on how to apply online, visit http://employment.umn.edu/ and search for Job Requisition No. 171635. Complete applications will be reviewed beginning September 2, 2011, and continue until the position is filled. The University of Minnesota is an equal opportunity educator and employer.

For focused results in accounting and tax research, start here.

While the world of accounting information continues [illegible] professional [illegible] a growing number of accounting students [illegible] reliable [illegible] for researching accounting trends, [illegible] issues, and history pertaining to the future of accounting.

ProQuest Accounting & Tax, our premier accounting research database, [illegible] research more effective [illegible] accounting [illegible] journals, reference reports, proceedings, dissertations, and more.

Researchers can [illegible] search over [illegible] publications with hundreds [illegible] full text, plus dissertations and working papers covering [illegible]

ABI/INFORM® subscribers can enhance their business database library with ProQuest Accounting & Tax, which includes [illegible] of additional full-text accounting [illegible] trade publications, dissertations and working papers not included in ABI/INFORM.

ProQuest Accounting & Tax includes:

- Highly ranked global and scholarly journals, including all journals from the American Accounting Association
- Accounting, tax management, and auditing trade publications from top publishers such as CCH, Aspen, Thomson Professional and Regulatory Services, Source Media, Tax Management Inc., and Euromoney
- AICPA publications
- International Federation of Accountants (IFAC) standards
- Tax law reviews from top institutions such as New York University, Stanford, Vanderbilt, and University of Virginia
- Tax law and financial law publications from noted sources such as the American Bar Association and American Law Institute
- Global accounting and tax journals from the UK, Canada, Australia, France, Spain, Ireland, and more
- Financial advising, governance, and risk management publications from Incisive Media, CFA Institute, and Association for Financial Counseling and Planning
- Accounting history from publications such as *The Accounting Historian's Journal*
- Over 1,500 accounting dissertations and more than 7,800 working papers from SSRN, plus select conference proceedings

ProQuest Accounting & Tax

Visit **www.proquest.com/go/accounting** today to learn more, and sign up for a free trial!

ProQuest®
Start here.